

7-23-2019

The Legacy of Hosting Mega-Sports Events: A Twitter Analysis of Leisure-Time Physical Activity Communication

Gabriela Baranowski Pinto

University of Connecticut - Storrs, gabriela.pinto@uconn.edu

Follow this and additional works at: <https://opencommons.uconn.edu/dissertations>

Recommended Citation

Baranowski Pinto, Gabriela, "The Legacy of Hosting Mega-Sports Events: A Twitter Analysis of Leisure-Time Physical Activity Communication" (2019). *Doctoral Dissertations*. 2222.
<https://opencommons.uconn.edu/dissertations/2222>

The Legacy of Hosting Mega-Sports Events:
A Twitter Analysis of Leisure-Time Physical Activity Communication
Gabriela Baranowski Pinto, PhD
University of Connecticut, 2019

Routine engagement in Leisure-Time Physical Activity (LTPA) is associated with long-term health, lower mortality, and higher quality of life. Inspiring participation in LTPA is one desired benefit of public investment in mega-events such as the Olympic Games. Whether such inspiration materializes—and, if so, for how long—was examined in three studies that investigated a perception-action process at the level of the interaction between real and virtual environments. The influence of the real environment of the Games in Rio de Janeiro in 2016 was assessed via the virtual environment instantiated in the Twitter social network. Emerging dynamic communication about LTPA in three epochs (before, during, and after the mega-event's occurrence) was assessed against pre-existing average weekly Twitter patterns not limited by topic. Study 1 compared Brazilian cities at different functional distances from the Games: the primary host, a football-only host, and a non-host. Study 2 examined the relevance of being close to or far from the host city's time zone, both for cities that had or had not previously hosted the Games. Study 3 targeted previous host cities (where English was the native language) over the last 60 years to examine persistence of engagement with the Games. Twitter activity comprised more than 1 million LTPA-focused posts collected during a span of 183 days from residents of 10 cities. Descriptive and inferential statistics were supplemented by Detrended Fluctuation Analysis, Fast Fourier Transform, Wavelet and Cross-wavelet Coherence Analysis, and Growth Curve Modeling. All analyses showed an influence of the Games on LTPA

Twitter activity in the targeted cities with the strongest effects being apparent in the cities of the host country and without systematic evidence that time zone or hosting history mattered. Average LTPA weekly patterns of behavior differed in each epoch relative to the Games, with none of the epochs strongly resembling the global patterns typical of activity not limited by topic. There was little evidence of what might be considered a legacy, whether in pre- and post-Games differences, recency of hosting, or even having hosted at all. Given evidence that tweets reflect actual behavior, the Olympic Games seem to provide the occasion that attunes the Twitter-user to the environmental affordances that encourage LTPA, but this effect does not seem to last long.

The Legacy of Hosting Mega-Sports Events:
A Twitter Analysis of Leisure-Time Physical Activity Communication

Gabriela Baranowski Pinto

Teacher Certification, Universidade Federal de Minas Gerais, 2006

BA, Universidade Federal de Minas Gerais, 2011

MA, Universidade Federal de Minas Gerais, 2009

A Dissertation
Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy
at the
The University of Connecticut

2019

Copyright by
Gabriela Baranowski Pinto

2019

APPROVAL PAGE

Doctor of Philosophy Dissertation

The Legacy of Hosting Mega-Sports Events:
A Twitter Analysis of Leisure-Time Physical Activity Communication

Presented by

Gabriela Baranowski Pinto, B.A., M.A.

Major Advisor _____
Claudia Carello

Associate Advisor _____
James Dixon

Associate Advisor _____
Steven Harrison

University of Connecticut
2019

Acknowledgements

The five years of engagement in this Program as a Ph.D. student presented me with constant challenges, which completely changed my life. The support of multiple people in a variety of ways was fundamental so I could finally go through that and fulfill all the requirements for getting the degree.

First of all, I would like to thank my dear advisor Dr. Claudia Carello who is an incredible mentor, a model of women scientists that I will never forget. She accepted me after she retired but it did not make any difference because she was always present guiding me and supporting me every step of the way. She was always generous in her criticism, turning it into a learning opportunity, and helping me to make sense of my dissertation research. I am truly honored to be her last advisee.

I would like to thank the associate advisors on my committee, Dr. J. Dixon and Dr. Steven Harrison, who spent a great deal of their time appreciating my research, asking questions, and looking for solutions to posed problems.

I recognize all members of the Center for the Ecological Study of Perception and Action (CESPA) and UConn professors for their patience and generosity in sharing all their knowledge, challenging me to be a better researcher. In particular, I remember Dr. Tehran Davis, who had a crucial role in my professional training for the Ph.D. while he was at UConn, helping me to adapt to the program, to understand perception phenomena, and to be involved in research. I also emphasize the goodwill of Dr. Michael Turvey always encouraging me to “make discoveries” and “work hard”, and for his more than useful comments on the presentation slides for my oral defense of this dissertation.

I thank all CESPA students (Cespians) for keeping such a good and friendly vibe, especially Adrian Frazier, not only for always being available for helping me to create and fix R codes, to think about statistical problems, or even for discussing my research with me, but also for keeping me company during many weekend and night times in the office.

I also would like to acknowledge Dr. Henry Harrison, an alumnus of CESPA, a wonderful programmer, who helped in the process of extracting the data for this research. I am grateful for the help of Renjie Chen from the Statistics Department with the wavelet analysis.

I highlight the importance of Dr. Alfredo Morales, who made himself available for answering questions about his reference paper, receiving me in Cambridge for a face-to-face discussion, and responding to many emails throughout the process.

I was lucky to encounter Dr. Dimitris Xygalatas from the Anthropology department, who accepted me as a Graduate Researcher in his lab and envisioned a series of projects related to my main interests. What I was able to learn from this experience has no limits and I will be eternally grateful for his confidence in me and for being an amazing mentor and friend.

My friends and family also have an important role in this victory. Even from far away they were always encouraging me to face my challenges, without pressuring me, demonstrating great pride for my accomplishments. I am truly grateful to my parents Luciene and Mauro, and to my Godmother Aunt Luiza, for believing in an education of quality and making real sacrifices to provide it to me. Also, thanks to my little brother Lucas, for staying at home helping out when I couldn't.

Vitor Profeta has a great role in this, because his dream about getting a Ph.D. abroad brought me to UConn. He was an amazing partner not only giving me moral and emotional support to accomplish my goals, but also being this amazing mentor who would always help me to solve my perception and action inquiries. It was a privilege to have you by my side.

Table of Contents

CHAPTER 1: Introduction	1
Social Networks	3
Sports Events and Communication about LTPA	4
Overall Goal	5
Three Studies	6
CHAPTER 2: General Methods	8
Research setting	8
Data Acquisition	8
Analysis Overview	9
CHAPTER 3: The Host City	14
Descriptive Statistics	15
DFA	16
FFT	17
Wavelet	17
CHAPTER 4: Study 1: Effect of Functional Distance from the Games	19
Descriptive Statistics	19
DFA	23
FFT	26
Wavelet and Cross-wavelet Coherence	27
Summary	31
CHAPTER 5: Study 2: Social Legacy vs. Time Zone Proximity	33
Descriptive Statistics	33
DFA	37
FFT	40
Wavelet and Cross-wavelet Coherence	41
Summary	45
CHAPTER 6: Study 3: Longevity of a Social Legacy	47
Descriptive Statistics	47
DFA	53
FFT	55
Wavelet and Cross-wavelet Coherence	57
Summary	60
CHAPTER 7: Comparisons of all cities	62
Slopes of Cumulative Sums during Three Epochs	62
Growth Curve Modeling	63
Summary	72

CHAPTER 8: General Discussion	73
Influence of the Games	73
Do the Games Leave a Legacy?	75
Concluding Remarks.....	76
References	80
Appendix 1	85
The Twitter metrics	85
Filters	85
Identifying LTPA hashtags	85
Identifying localities associated with targeted cities	85
Appendix 2	86
Detailed Patterns	86
Study 1	87
Study 2	88
Study 3	89

CHAPTER 1

Introduction

When governments decide to bid to host large-scale international sports events such as the Olympics, they typically tout the legacy of that event for their people. A legacy is usually understood as some kind of economic, infrastructural, social, cultural, environmental, or any other kind of benefit left in the present or that will remain in the future (Mazo, Rolim, & Dacosta, 2008). So-called mega-events are tremendously expensive, averaging more than \$10 billion since 2000 (McBride, 2018) so it is hoped that the benefit goes beyond an increase in civic pride (Groothuis & Rotthoff, 2016), and reaches the legacy realm.

It is a recognized fact that much of the cost for the event goes to improved infrastructure, an obvious, tangible legacy. Subways have to be upgraded to handle the influx of fans, housing has to be built to accommodate the visiting athletes, and these structures will remain for the benefit of the host population, as a societal benefit. Although, these are positive outcomes of this type of event, they are often undermined by subsequent maintenance and conversion costs. Also, they are accompanied by an exaggeration in the predictions, especially with respect to host cities of countries under development (Swinnen & Vandermoortele, 2008).

Economics-related legacies of mega-events are more evident in the news highlighted by the media and by the event organizers than other types of legacies (Hall & Hodges, 1996). But might the economic cost be offset by less tangible benefits? Invisible social legacies may arise in the domains of health, professional training, or social networks constituted by the event. Improvement in health, in particular, may derive from increasingly active lifestyles inspired by media attention on physical activity during the event (Minnaert, 2012). It is common to find reports of increased participation in specific Olympic sports activities (e.g., swimming, running, volleyball) and general healthful activities (e.g., aerobics, yoga, home gym exercise) in the immediate aftermath of a mega-event (Sporting Goods Manufacturers Association,

2012). However, there are inconsistencies in this matter. Many studies failed to show change in behavior related to so-called Leisure-Time Physical Activities (LTPA), even in developed countries (Veal, Tooheyb, & Frawley, 2012; Weed et al., 2009). The possible legacy for LTPA provided the focus for the present dissertation.

One way to frame the issue is with respect to so-called *objective necessities* (Bourdieu, 1984), voluntary activities that have, in a very real sense, become obligatory. As an example, many people consider checking the internet to be no more discretionary than biological demands such as eating, sleeping, and so on. Under appropriate circumstances that inspire a legacy, it might be expected that LTPA should become such an objective necessity, too.

Typical indicators of LTPA are individual and local—for example, a person in the host city joins a health club. But these may miss the possibility of broad-based social interactions as a driving force of increased activity. The 21st century has been marked by the rapid emergence and importance of social networks, on-line amalgamations of the activity of people who become connected through the sharing of facts, reactions, feelings, and opinions. Events occur in natural environments and people communicate about them, directly and indirectly, in the virtual environment of the network. Networks such as Facebook, Instagram, and Twitter provide a potentially rich setting in which to study interpersonal interactions on a large scale. We can ask whether an event perceived in the physical environment affects actions presented in the virtual environment and vice versa. In order to address this question, an event (the Olympic Games) and an action (Twitter communication about LTPA) were selected for investigation. Specifically, we asked how communication about LTPA emerged in social media as a function of the imminence, occurrence, and aftermath of the Rio Olympics. To that end, the virtual environment of social networks as a medium and LTPA as the focal action will be motivated.

Social Networks

Although the complexity and fundamental essence of human-machine networks still need to be fully understood (Pentland, 2014), topic-specific data can be extracted that provide rich information, especially if they are time-sensitive and geographically specific (Bollen, Mao, & Zeng, 2011). A complex systems perspective on the dynamics of large groups suggests that people's activities (including physical activity) are responsible for the emergence of the large-scale patterns that define them (França et al., 2016; cf. Bar Yam, 1997). In the case of social networks, it is possible to measure the emergence of patterns of agreement in communication, giving an indication of the extent to which private activity and sentiments cohere in an interpersonal fashion. The fact that users share topics through posts, comment on them, or express their appreciation, provides insight about points of agreement on a specific topic such as LTPA.

There are several advantages to the use of the Twitter archive in research. It is a direct record of spontaneous activity, emotions, and thoughts. Since it provides time-lines of daily communication about a diversity of topics, we can discern the dynamics of how communication is coordinated among members of the same social interface and how the dynamics change as a function of societal events. Not requiring participants' self-reports yields observations that are neither voluntarily nor mistakenly disconnected from reality. Also, it avoids imposing or changing the context in which the individual is immersed (as would be the case with experimental manipulations of the environment or the use of observers in the field). In this setting, communication is the medium that makes it possible to quickly grasp tendencies at societal, community, and individual levels.

Hypotheses were tested on the Twitter social platform in which users post short status messages called tweets (via internet or mobile phones) as a microblog and as a means to connect with other members by following them as a social network. Users are directly informed about their posts, which—

through retweets, for example—can spread through the network. Twitter’s main purpose is spreading news about offline public events or minor daily activities (Thelwall, Buckley, & Paltoglou, 2011), providing insights into public emotions and opinions about a wide range of topics. It has the capacity to influence a variety of broad domains, for example, governments (Pentland, 2014), the stock market (Bollen et al., 2011), and public sentiment (Tan et al., 2014). The focus of this research was mostly on the microblog properties of the network.

Sports Events and Communication about LTPA

A good deal of on-line discussion is devoted to sports, accounting for an estimated 13% of social media posts of adults in the United States as of June 2016 (Statista, 2018). This presence increases around the time of a big event. For example, the 2016 Super Bowl elicited seven times the number of tweets as the next most popular football game (Statista, 2018). At issue for present purposes is whether such an increase in on-line activity indexes a social legacy of stimulating the practice of LTPA. This issue may be addressed using social networks as an instrument to access LTPA indicators. Spontaneous media activities, generated not only locally but also worldwide, will increase exposure to sports topics on a daily basis. This might influence LTPA behavior on a daily basis, with such changes expected to be reflected, in turn, in generated Twitter communication related to the event and potentially to LTPA. A change in the timescales of occurrence of communication about these activities after the event could provide evidence of a change in what activities users now consider to be objective necessities (Bourdieu, 1984). Interpersonal consistency in communication could be a legitimate expression of human collective entrainment with events in the environment (cf. the analysis of Twitter activity about the Presidential debates; Fusaroli et al., 2015).

A social media network such as Twitter could be used as a tool to provide a better picture of existing constraints on the distance, timing, and connectivity of communication of large populations

experiencing those events. Put another way, Twitter provides indications of how massive, unfolding real-world events drive and resonate with human social behavior in a precise, temporal dynamics (Fusaroli et al., 2015).

Overall Goal

This research investigated how emerging dynamic communication about LTPA in the Twitter virtual social environment was influenced by a particular large-scale event—the Olympic Games—before, during, and after its occurrence. Temporal patterns of LTPA communication in targeted cities—the host city, another city in the host nation, a non-host Brazilian city, international cities, and former host cities—were assessed. As a standard of comparison, we will refer to results from previous research on weekly Twitter patterns in urban areas across the globe based on tweets undifferentiated by topic (Morales, Vavilala, Benito, & Bar-Yam, 2017). Figure 1 shows Twitter activity. It reveals three distinct patterns across the globe, represented here by Rio de Janeiro, Sydney, and London. They are distinguished by different shapes of the peaks within each weekday (a graduated ramp-up to a prominent peak, a double peak with one much more prominent, double-peak that are more equal in amplitude). Regardless of pattern, however, Fourier transforms of those time series revealed that three periodicities stood out: 8, 12, and 24 hr. Questions to be addressed here included whether such patterns are affected by: (i) focusing on a single topic, LTPA; (ii) the temporal proximity of the Olympics; or (iii) the spatial or functional proximity of the targeted city to the host city.

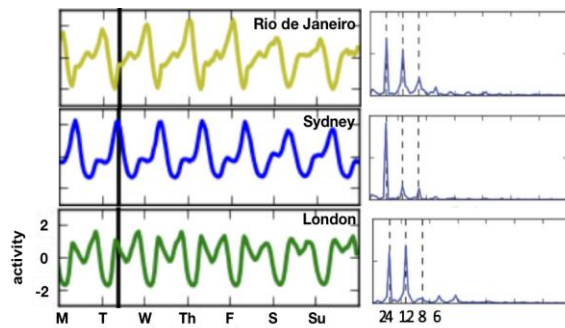


Figure 1. (left) Three patterns of normalized average weekly Twitter activity identified in Morales et al. (2017) are represented by three cities. (right) Corresponding Fourier transforms for those same cities show three significant components at 24, 12 and 8 hr.

It would not be surprising that a big event affects Twitter communication about the event itself: people chatting about particular sports or medals or national results. But the present research concerns personal physical activity. Does the event inspire communication about the individuals' own LTPA? It was hypothesized that dynamic communication would change for all targeted cities, indexed by an increase in the number of tweets about LTPA, with an accentuated effect for the host city right before and during the Games. After the Games, this activity was expected to gradually relax to lower levels. Additionally, the collective dynamic that emerges in communication will be characterized as a further exploration of changes that might provide more subtle indications of a kind of legacy. Whether relaxation levels, collective dynamics, and so on differ between the host city and the targeted cities are all relevant for the notion of a legacy. Broadly, the greater involvement of a particular city in the Olympic Games, the stronger the consequences for LTPA are expected to be.

Three Studies

In order to accomplish these goals, three studies were developed in this dissertation. Study 1 compared Brazilian cities at different functional distances from the Games: the primary host (Rio de Janeiro), a football-only host (Belo Horizonte), and a non-host (Porto Alegre). Study 2 examined the relevance of being close to (U.S. Eastern Daylight Time) or far from (Australian Eastern Standard Time) the host city's time zone (Brasilia Time) for cities that had previously hosted the Games (Atlanta and Sydney) or had not hosted (Philadelphia and Brisbane). Study 3 targeted previous host cities (from

Melbourne in 1956 to London in 2012) in order to examine whether time since hosting the Games would matter to the degree of social media engagement with the Games. In order to ensure that tweets were really about LTPA, former host cities were limited to those where English was the native language.

CHAPTER 2

General Methods

Research setting

This dissertation was inspired by the approach of Morales et al. (2017) who examined geo-located tweets (i.e., identified by longitude and latitude) for all topics discussed in the network across the globe (some 500 million tweets per day). We targeted tweets based on users' profile location reports (i.e., home city) with the focus on the global patterns of Twitter communication about a particular topic, namely, LTPA.

Data Acquisition

The Olympic Games are one of the most popular mega sports events in the world, involving in excess of 100 countries. The most recent iteration, hosted by Rio de Janeiro, Brazil in 2016, was chosen as the test setting to anchor the data acquisition timeframe. Six Twitter metrics were obtained: tweets, retweets, chatters, number of users, average number of hashtags per tweet, and average number of words per tweet¹. The filtering of the metrics encompassed a "scraping period" from June 1 to November 30, 2016. This covers 65 days prior to the Olympic Games, the 17 days that constitute the Games (August 5-21), and 101 days after the Games. The data were downloaded in real time from the Twitter streaming Application Program Interface (API) as implemented by Tweepy (a Python interface) to the limit of 1% of the total number of tweets being generated, which on average is 60 tweets/s. The data selection occurred at the same moment the tweets were generated and lasted the same number of days as the total scraping period. The tweets downloaded from the API were focused on hashtags related to LTPA and filtered for target city. (See Appendix 1 for how hashtags were identified and locations were filtered.)

¹ A description of these metrics can be found in the Appendix 1. Given the similarity of the patterns for these metrics, not all analyses will address all of them.

Of course, cities differ in how much they use Twitter. Most obviously, activity is affected by the size of the population as well as the age of its residents—more people, especially young people, is associated with more tweets. And there are, no doubt, a variety of other socio-cultural factors (Kulshrestha et al., 2012 Speed, 2014). Comparing cities that differed in absolute numbers of tweets was addressed by normalizing each city's data relative to its own Twitter habit, here represented by the total number of tweets over the entire scraping period. Additionally, given the strong influence of the day-night cycle on Twitter patterns, comparisons were further facilitated by rendering activity in each city's time zone.

Analysis Overview

Descriptive statistics. In order to better visualize the characteristics of each Twitter metric in every targeted locality, a variety of descriptive statistics was explored. Given the duration of the Olympic Games, multiple windows of 17 days were created. Descriptive statistics were computed for the three pre-Games windows (starting at 51, 34, and 17 days before the Games), the Olympic Games window itself, and for five post-Games windows (ending 17, 34, 51, 68, and 85 days after the Games). The sum, maximum, mean, median, and standard deviation (SD) of the number of tweets produced per hour in each window, as well as the cumulative sum of tweets throughout the entire scraping period of data collection were obtained for every city (see Appendix 2). The weekly averages per hour were also obtained for the pre-, during, and post-Games epochs to assess potential periodicities. Finally, the frequency distributions of the same variables per location in every window of interest were generated.

Measuring this type of emergent behavior in social networks is a challenge because of its complex structure. Count data, for example, is often not normally distributed, being extremely skewed and following power laws. This characteristic poses a challenge for analysis, which in this dissertation was overcome by using techniques that are more sensitive to the non-linearity underlying the data structure from social networks. These techniques are going to focus primarily on long-term correlations and

oscillatory behavior, but also on longitudinal evolution of the data set being assessed. These kinds of techniques have been seen before in work on self-organization of social media political activism (Aguilera et al., 2013), responses to extreme events in social media (Gao & Liu, 2015), Twitter activity at a global level (Adnan et al., 2014), and establishing connections between social media sentiments and the stock market (Xu et al., 2017).

We predicted that Brazilian cities would present relative higher levels of tweet production than the other targeted cities and that the period of the Games would have considerably higher means, median, and maximum values.

Detrended Fluctuation Analysis (DFA). DFA was performed in the three studies in order to identify and describe the long-term correlation behavior embedded in the time series of each locality. The presence of long-term correlations is also referred to in the text as dependency. For the Twitter metrics analyzed here, the DFA generates the Hurst exponent, a scaling index of the structure of noise across time scales. This exponent formalization corresponds to $\beta = 2H - 1$ for non-stationary cases. At one extreme, a Hurst exponent (H) of 0.5 indicates random patterns (white noise) with no evident dependency or correlation and average deviation from the mean equal to zero. At the other extreme, a Hurst exponent of 1 indicates some relationship (pink noise) between the observed patterns at different time scales, with dependency tendencies of any value to be succeeded by something similar. In other words, it means "... high values tend to be followed by more high values" (Brown & Liebovitch, 2010, p.70).

DFA for all the targeted localities was performed in two phases: (1) for the different length time series of pre-, during, and post-Games (65, 17, 101 days respectively); and (2) for each of the 17-day windows (three pre-Games, one during the Games, and five post-Games. Since, we did not have much reference of possible outcomes of this analysis, we decided not establish predictions about it.

Timescales found in the patterns of tweets of targeted cities. If hosting a mega event such as the Olympic Games changes the dynamics of communication on LTPA, this should be reflected in the timescales of communication as revealed by a Fast Fourier Transform (FFT) and Wavelet analysis. Morales et al. (2017), for example, identified three prominent timescales for worldwide Twitter network activity not limited by topic: 24, 12, and 8-hr frequencies. It is reasonable to expect that this signature of daily collective activity would be altered during the Games, perhaps especially in local synchronies that emerge in the host city. For example, the emergence of shorter timescales might provide evidence of a social legacy reflecting more interweaving of LTPA in people's lives. Relatedly, it was expected that a daily timescale of communication activity would be more appropriate than a weekly timescale, considering that people would bring LTPA to their daily lives and not, as the literature typically suggests, only to weekends when the amount of free time is considerably higher (Godbey, 2008). These kinds of changes would speak to the emergence of LTPA as objective necessities (Bourdieu, 1984).

Fast Fourier Transform (FFT). The power spectral analysis was used to identify significant frequencies in each of the 17-day windows. The average week for each window was considered so as to allow comparisons with the all-topic Twitter network activity presented by Morales et al. (2017). The FFT analysis for all the windows set the taper = 0.1 - 0.5 and used the Daniell kernel = 1. Periodograms were plotted after a Y-axis scale normalization was performed. Power values of each time series were divided by the maximum power level and varied between 0 and 1, with values closer to 1 being more powerful.

The identification of relevant frequencies in the spectrum was based on a peak-picking technique. After plotting the periodogram, a threshold of 0.2 was established and the peaks above it selected as the most prominent frequencies that contributed to the final signal (Figure 2). The findings to be reported are based on the period with respect to the highlighted frequencies, which is a product of 1 divided by the identified frequency.

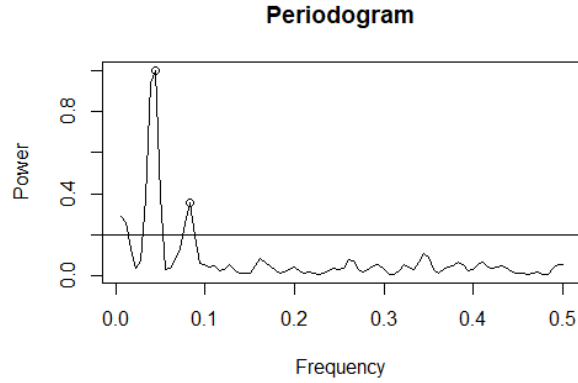


Figure 2. Example of the procedure to highlight in the periodogram the relevant frequencies generated by FFT. The peaks above the threshold = 0.2 (black horizontal line) are indicated.

Wavelet and Cross-wavelet Coherence. Because the FFT emphasizes only those frequencies that contribute the most to the signal of interest, it may miss details about less prominent frequencies. Consequently, a wavelet analysis was performed for each time series of each Twitter metric for every targeted locality. The power spectrum scale was based in quantiles, and the mother wavelet used was the Morlet. *P*-values of the variance were computed using white noise as the method of generation of surrogates to represent the null hypothesis (also not typical for the FFT analysis). Other relevant specifications are: a Loess detrending of 0.75, time resolution $dt = 1$, dominant frequency resolution $Df = 1/20$, lower Fourier period = $2 \cdot dt$, upper Fourier period = floor of $1/3$ of time series length $\cdot dt$.

In addition, the cross-wavelet coherence between Rio de Janeiro and each targeted city was calculated, normalized by variance across time. This analysis addresses the cross-spectrum time-localized coherence and relative phase of the paired time series at different frequency ranges (i.e., time scales) across the 183 days of Twitter activity in each city (Romero et al., 2018), through 10 Monte Carlo randomizations. In this analysis the mother wavelet used was also the Morlet. Visual inspection of the cross-wavelet coherence plots suggested that the analysis should be conducted for the periods 12 ± 1 hr, 24 ± 1 hr, and 256 ± 1 hr (i.e., band-width 1 considered one hour before and after each targeted period). The significance of the results was calculated based on the ratio between the outputs *rsq* and *signif*, which are the matrix of wavelet coherence and the matrix of percentiles of significant levels of wavelet coherence based on Monte Carlo autoregressive time series.

Growth Curve Modeling (GCM). Given that this research dealt with retrospective datasets facing longitudinal changes, a particularly useful procedure for capturing the development of a dependent variable as a function of time is multilevel Growth Curve Modeling, or GCM. The frequency of tweets posted in the network throughout the scraping period constituted the dependent variable. In the present case, the total number of tweets was accessed repeatedly in a continuous fixed and identical schedule for all targeted cities, meaning the frequency of tweets of the entire network was sampled every hour. Essentially, since the hours can be clustered in multiple ways, several timescales were explored (hourly, daily, weekly, and 17 days). All cities targeted in this dissertation were used together in the GCM. The predictors tested in the models for this analysis were essentially categorical and addressed the differences in tweet production as the dependent variable for the conditions of 1) being a Brazilian city vs. not, 2) being proximate to the time zone of the Games vs. not, and 3) being a previous host of the Games vs. not.

CHAPTER 3

The Host City

As the contemporaneous Olympic host during the scraping period, Rio de Janeiro provides the anchoring for comparisons in all three studies. This section is devoted entirely to the particulars of this city before the relevance of functional distance, time zone, and previous hosting status can be evaluated.

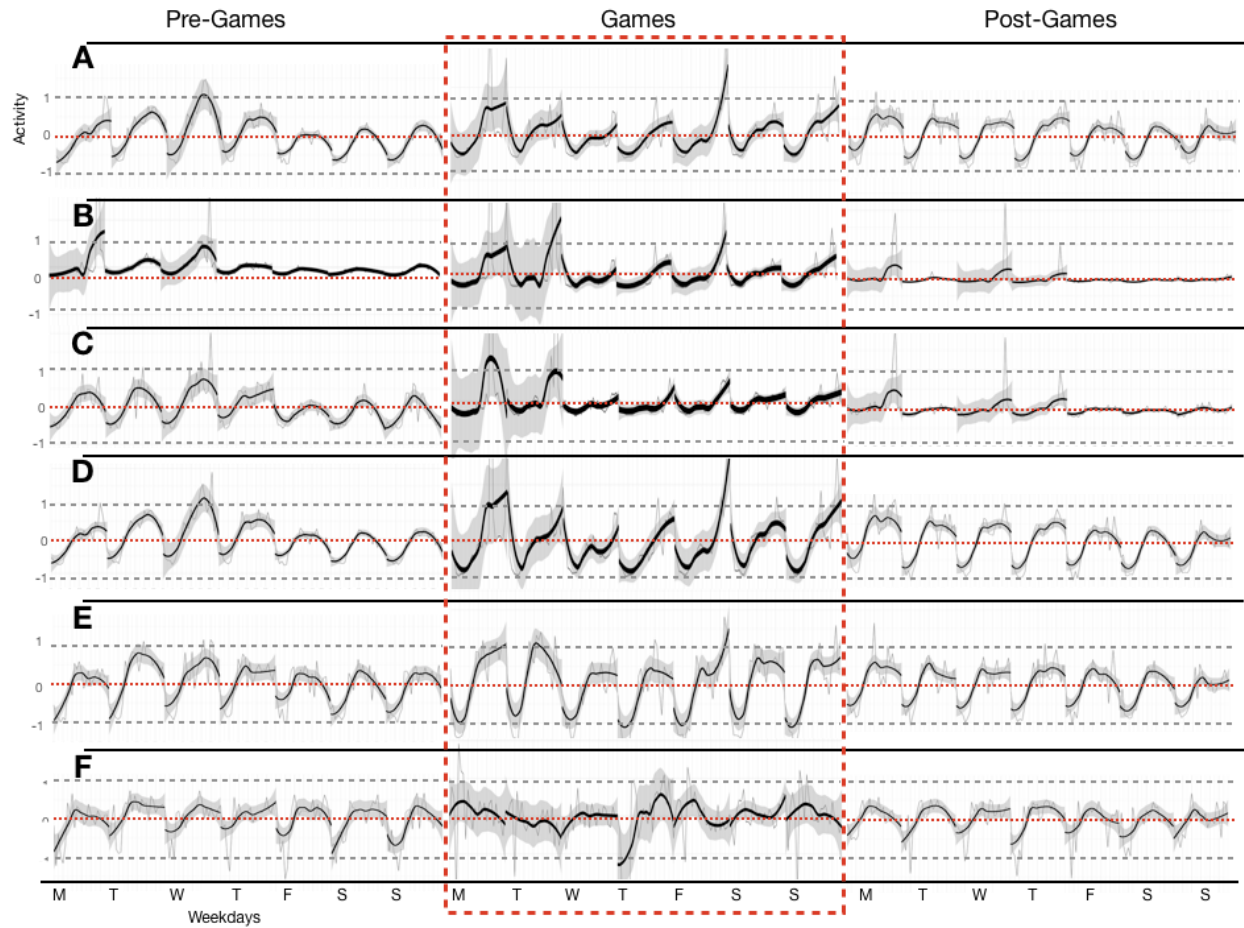


Figure 3. LTPA-related tweets (A), retweets (B), chatters (C), users (D), average number of hashtags (E), and average number of words (F) for Rio de Janeiro during three epochs: (left) pre-Games, (middle) during Games, and (right) post-Games. Activity for each metric is averaged over corresponding hours during each day of the week throughout the scraping period. The weekday average number of tweets (m) is shown at the top of each weekday.

Descriptive Statistics

Weekly average LTPA Twitter activity in Rio de Janeiro is shown in Figure 3. Inspection of pre- and post-Games patterns reveals typical Twitter behavior reflecting day-night cycles and sleep time: a decrease in activity for all Twitter metrics around 12 pm and in early mornings, putting the normalized data in the negative region, with increased activity as the morning progresses, reaching a peak around 8 am that lasts until around 8 pm, when system activity starts to decrease. This pattern is characteristic of any locality around the world, no matter the cultural characteristics. During the Games, however, the classic sleep time zones vanish into a flat pattern for almost all metrics, with the exception of average number of hashtags, which maintains the characteristic oscillatory behavior. In addition, note the difference between this LTPA tweet activity and the pattern of all-topic tweet production (Figure 1) obtained for this same city three years earlier (cf. Morales et al., 2017). It is hard to discern the graduated pattern of “two small peaks of activity in the morning and a large peak in the evening” that is typical of unrestricted-topic tweets in multiple South American cities.

Overall, normalized activity varies within ± 1 SD from the mean. During the Games, however, variability is not as stable—moments in the week show much more variability, going well beyond ± 1 SD, sometimes surpassing 5 SDs from the mean. It is also notable that the pre-Games period is characterized by decreased variability of tweets, retweets, chatters, and users during the weekends but this pattern does not persist during the Games nor does it return after the Games.

Figure 4 shows the proportional daily Twitter activity of tweets, retweets, and chatters produced throughout the scraping period. In general, individual Tweets comprise the bulk of activity. It is apparent, however, that during the Games there is a relative increase in the *reproduction* of information as the proportion of retweets increases. The elevation of retweets as a share of overall activity seems to persist after the event finishes.

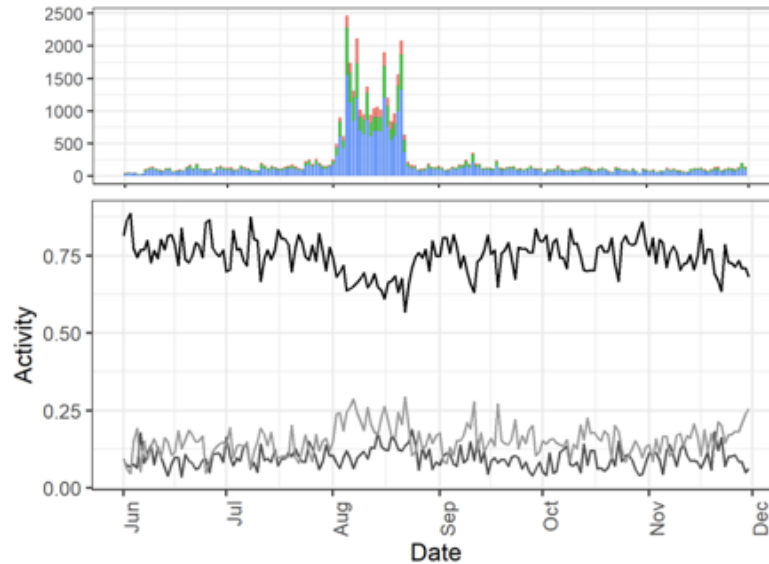


Figure 4. (top) Histograms of daily LTPA-related tweets (blue), retweets (green), and chatters (pink) over the entire period of data collection for Rio de Janeiro. (bottom) Time series of proportional daily LTPA-related Twitter activity metrics in Rio de Janeiro: Number of tweets (black), retweets (light grey), and chatters (medium grey). Vertical dashed lines delimit the occurrence of the Games.

DFA

Pre-, during, and post-Games. The value of H for the host city generally exceeded .75 with the exception of the pre- and post- epochs for number of words used, which were nonetheless greater than .60. In short, DFA revealed persistence for all metrics in all time periods. Interestingly, although the value of H was the greatest during the Games for number of hashtags and number of words, this same period yielded the smallest H for tweets, retweets, chatters, and users.

17-day windows. The results for the finer partitioning largely echo the foregoing. For tweets, retweets, chatters, and users, dependency is greatest further away from the Games, both before and after the event. The dynamics move in a randomized direction just prior to the Games after which there is a kind of recovery—for the last post-Games window we examined, H returned to the highest level that was seen in the first window we examined. For number of hashtags and number of words, the peak during the Games was preceded and followed by a steep decline.

FFT

The frequency analysis of Rio de Janeiro yields the periods 22.5 or 25.7-hr in all windows, approximating a 24-hr daily pattern as expected in Twitter data. The 12-hr period is also above threshold in all but window 3, the last time series before the Games began. Windows 3 and 5 (immediately before and after the Games) yielded relatively low frequencies of 60 and 90 hr. These frequencies vanished during the Games, when the periods 9 and 6.4 hr emerged—the shortest periods encountered in Rio’s FFT analysis. In summary, while the system we are evaluating becomes low frequency at the boundaries of the Games, fast frequencies appear during the event, indicating that people might be engaging more often in tweeting about LTPA (Figure 5).

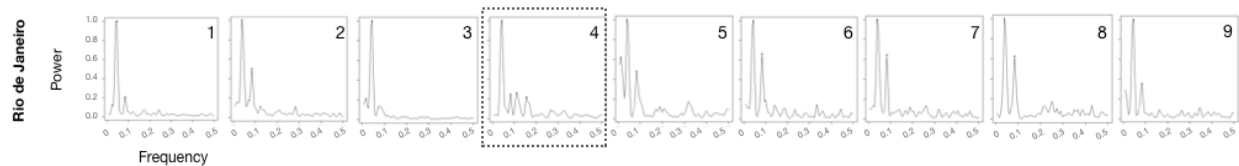


Figure 5. Periodograms generated from the FFT for 17-day windows of tweet activity in Rio de Janeiro. (1-3) Pre-Games windows corresponding to 51, 34, and 17 days before the Games; (4) the Olympic Games window; (5-9) post-Games windows corresponding to 17, 34, 51, 68, and 85 days after the Games.

Wavelet

The Wavelet Power Analysis performed on the entire Rio de Janeiro time series allows an evaluation of the influence of the offline event on the spectrum magnitude (Figures 6). This analysis echoes the appearance of 24 and 12-hr main frequencies, albeit significantly so only during the time around the Games. Finer detail is available in the sequence of 17-day windows, however, where we see the predominance of the 24-hr period in most of the windows (Figure 7). Except for windows 3, 4, and 5 (corresponding to right before, during, and after the Games), this was also accompanied by a high power. Additionally, although the 12-hr period seems to be an important frequency, it does not persist through the entire time series.

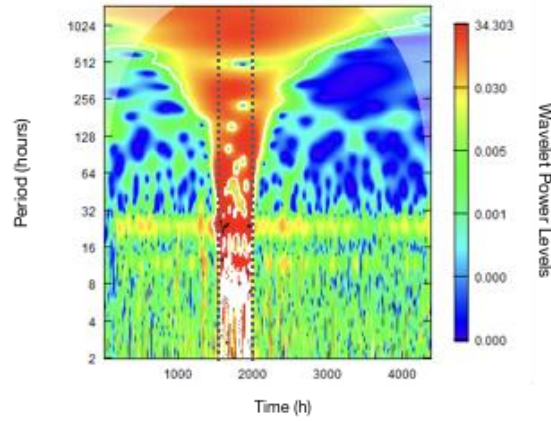


Figure 6. The Period \times Time Wavelet Power Spectrum generated for the entire period of tweet activity collected in Rio de Janeiro (a total of 4391 hr). The Y-axis shows the candidate frequencies; the X-axis shows the time from the beginning to the end of the time series; color indicates the power level. Vertical dashed lines delimit the occurrence of the Games (between 1560 and 1967 hr). (Regions that are significantly different from white noise are outlined in white.)

Significant fast frequencies are also apparent, though less so in Windows 3 and 4. The occurrences of high magnitude fast frequencies are much more concentrated at certain times and are also persistent across different moments. That is, many frequencies have high power and contribute to the final signal. The long vertical red patch in window 3 reveals considerable power across many frequencies around 100 to 50 hr before the Games. Statistically significant power is also apparent around 64 hr in all windows but 7 and 8, but its contribution is brief in every window and at a substantially lower power level, suggesting its contribution to the final signal of the time series does not have much relevance.

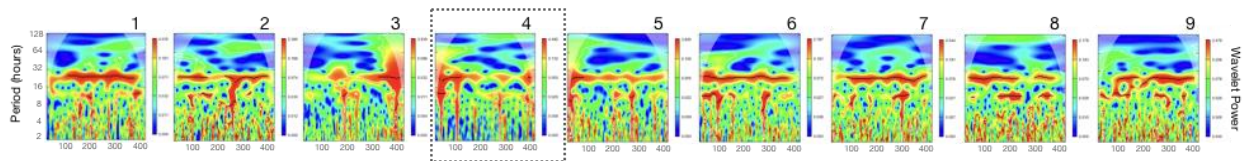


Figure 7. Period \times Time Wavelet Power Spectrums generated for successive 17-day windows of tweet activity in Rio de Janeiro. (1-3) Pre-Games windows corresponding to 51, 34, and 17 days before the Games; (4) the Olympic Games window; (5-9) post-Games windows corresponding to 17, 34, 51, 68, and 85 days after the Games. The Y-axis shows the candidate frequencies; the X-axis shows the time from the beginning to the end of each window (Regions that are significantly different from white noise are outlined in white.)

CHAPTER 4

Study 1: Effect of Functional Distance from the Games

Rio de Janeiro was the principal host city, featuring wide-ranging events including track & field, swimming, cycling, boxing, weight-lifting, sailing, basketball, volleyball, and football (soccer). While several other cities in Brazil hosted football, other cities did not host any events. In other words, Brazilian cities differed with respect to what we might call their *functional distance* from the Olympics. The influence of functional distance on the patterns of communication was examined by comparing LTPA Twitter activity in Rio to the football host Belo Horizonte, and to the non-host Porto Alegre. These cities were chosen based on their metropolitan population size (PS), with Social Vulnerability (SVI) and Human Development Indexes (HDI) that were close to the standards for Rio de Janeiro (PS = 12,389,775; SVI = 0.290; HDI = 0.799). Belo Horizonte (PS = 5,742,260; SVI = 0.276; HDI = 0.810) was a football host; Porto Alegre (PS = 3,894,232, SVI = 0.249, HDI = 0.805), did not host any events. All three cities belong to the South-Southeast region of Brazil.

Descriptive Statistics

The analysis considered 31,605 tweets in Rio de Janeiro; 9,315 in Belo Horizonte; and 3,746 in Porto Alegre. In terms of absolute numbers, Rio tweets more than the other cities in all epochs (Table 1). Total LTPA tweets during the scraping period as a percentage of population (considering each municipality plus its metropolitan area) were 0.24 for Rio, 0.16 for Belo Horizonte, and 0.10 for Porto Alegre. These percentages indicate that differences in Twitter use among cities are not simply due to population differences. As noted, usage may reflect socio-cultural factors. In Brazil, for example, the ease of accessing wireless networks may influence local tendencies. This fact reinforces the need to normalize each city's data relative to its own Twitter habit.

Table 1
Mean Number of Tweets per Hour per Day (with SD) during Three Epochs Relative to the Olympic Games for Three Brazilian Cities.

City	Epoch	Tweets
Rio de Janeiro	51-34 days prior	3.65 (3.23)
	17 days during	37.11 (46.81)
	34-51 days after	3.48 (2.91)
Belo Horizonte	51-34 days prior	1.397 (1.313)
	17 days during	7.916 (9.839)
	34-51 days after	1.512 (1.069)
Porto Alegre	51-34 days prior	0.358 (0.649)
	17 days during	3.754 (5.930)
	34-51 days after	0.414 (0.730)

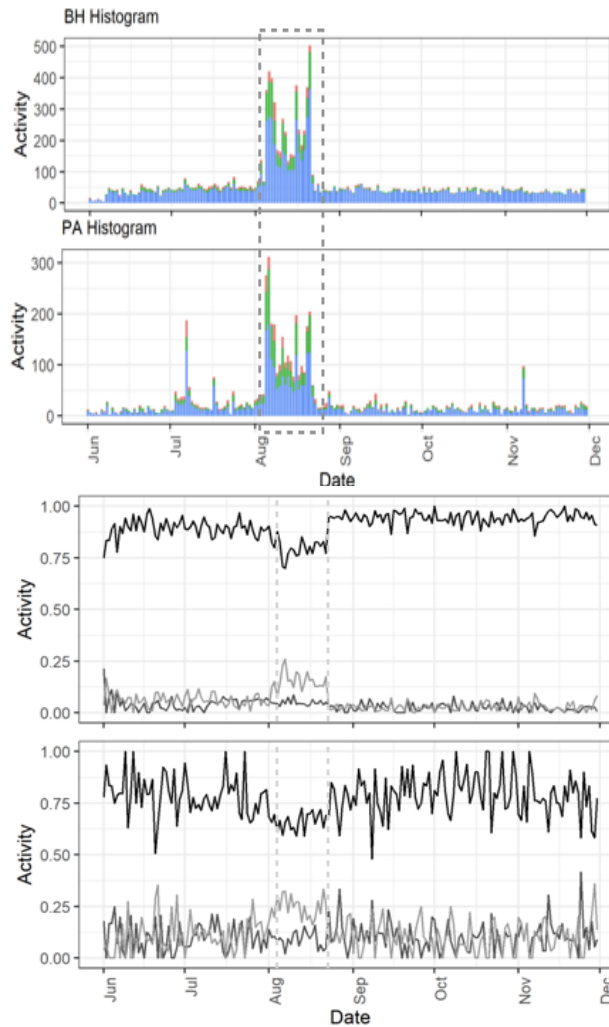


Figure 8. Histograms of daily LTPA-related tweets (blue), retweets (green), and chatters (pink) over the entire period of data collection for Belo Horizonte (BH), and Porto Alegre (PA).

Figure 9. Time series of proportional daily LTPA-related Twitter activity metrics in Belo Horizonte (top) and Porto Alegre (bottom): Number of tweets (black), retweets (light grey), and chatters (medium grey). Vertical dashed lines delimit the occurrence of the Games.

Twitter activity increased in all cities during the Games, with the increase being sharpest for Rio de Janeiro (details in Appendix, Figure A2-1). While the average number of hashtags per tweet also follows

this basic trajectory, it is distinct from the pattern for tweets, retweets, chatters and users in that, prior to the Games, Rio de Janeiro's rate is notably higher than Belo Horizonte's which is higher than Porto Alegre's. Moreover, the sharper increase for Rio during the Games is not as dramatic as for the other metrics. The differences among the cities continued immediately after the Games but began to diminish near the end of the scraping period (Figure A2-1). Exceptions to the typical trajectory were found in the maximum number of words per tweet and their standard deviation, both of which decreased during the Games.

The increase in tweet activity in Figure 8 coincides with the period of the Olympic Games. Figure 9 shows that, as for the host city, individual Tweets comprise the bulk of activity in both cities and that the relative contribution of retweets increased during the Games. Unlike the host city, however, after the event the retweet levels in both cities returned to their pre-Games level.

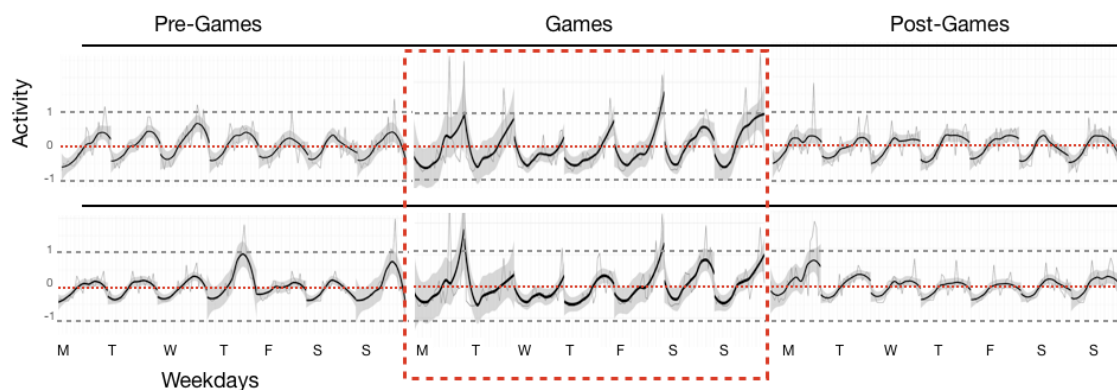


Figure 10. Weekly LTPA-related tweets for Belo Horizonte (top) and Porto Alegre (bottom). Pre-Games (left), Games (middle), and post-Games (right). Activity is averaged over corresponding hours during each day of the week throughout the corresponding scraping period. Data were normalized (z-scores) for the entire scraping period. Red dashed lines indicate the mean; grey dashed lines indicate ± 1 SD.

The weekly patterns of tweet production reveal that before and after the Games both cities' activity is within ± 1 SD from the mean; during the Games, in contrast, variability is magnified (Figure 10). Unlike Rio, however, there is less variability in classic sleep time zones do not vanish during the Games. There is also the suggestion, at least on some weekdays before and after the Games, of the graduated double-peak pattern reported in Morales et al. (2017) for all-topic tweets in multiple South American cities. It

might indicate that the fashion of tweeting about LTPA topics shares the same local synchronies evident in the patterns of activity about other topics, at least for cities at some remove from primary hosting duties. Both cities' weekends activity, especially on Sundays, seem to present more variability away from the mean.

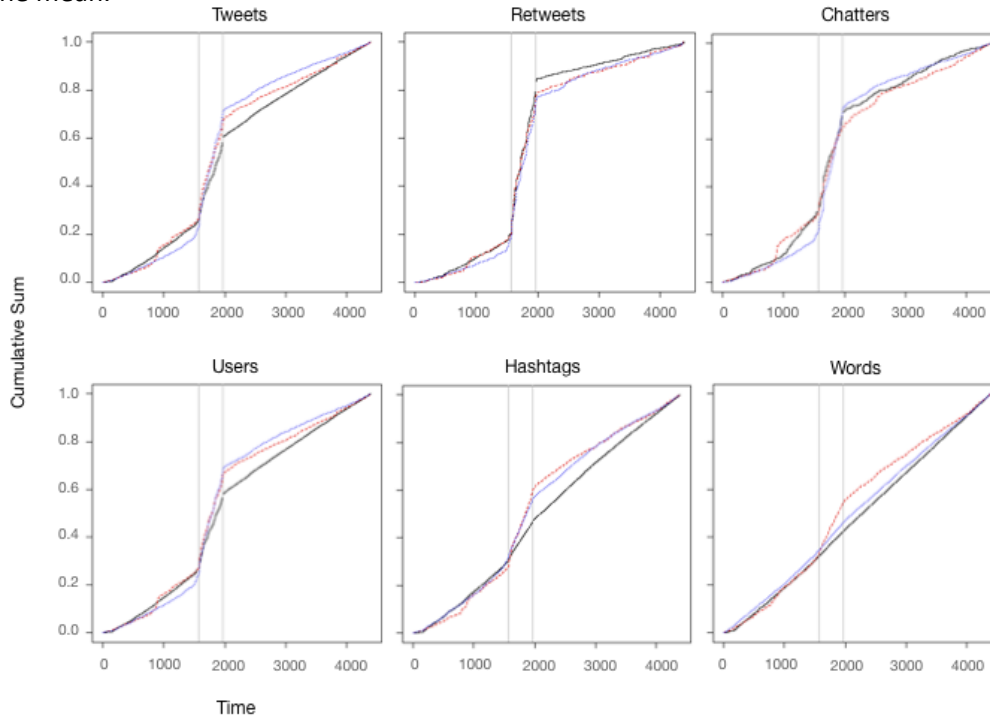


Figure 11. Cumulative Sums for tweets, retweets, chatters, users, average number of hashtags, and average number of words over the entire period of data collection for Rio de Janeiro (blue dotted line), Belo Horizonte (black line), Porto Alegre (red dashed line). Data are normalized by each city's total number of tweets during the scraping period. The occurrence of the Games is bracketed by the dashed vertical lines.

Figure 11 shows the cumulative sums of the Twitter metrics. To facilitate comparisons across cities that differ in their general use of social media (e.g., due to population, accessibility of wireless networks, local habits), each metric was scaled by the total activity during the scraping period so that they all run from 0 to 1. Inspection of the plots reveals exponential growth during the Games for all observed metrics, in all three cities, with two exceptions: Linear, non-exponential growth characterizes the average number of words per post for Rio and Belo Horizonte, and the average number of hashtags for Belo Horizonte. Otherwise, the increase in the production of tweets has very similar slopes for the three cities, with the steepest slopes occurring during the Games. On the other hand, the pre- and post-

Games regions of the cumulative sum have decreased slope inclinations with shallower patterns that indicate a drastic change in the rate of tweet production. This kind of growth is common in Twitter data due to the contagious nature of the spread of information, which generates S-shaped curves (Christakis & Fowler, 2010).

DFA

Figure 12 presents results of the DFA for the entire scraping period before the Games, the 17 days of the Games, and the entire scraping period after the Games. Values of H are between .5 and 1.0 for all metrics throughout the scraping period, indicating persistence. And although the degree of persistence in the three cities varied during the pre- and post-Games periods, it is notable that the values of H converged during the Olympic Games. In short, it appears that the occurrence of the mega-event impacted the social network's manner of talking about LTPA. A few city-specific patterns deserve comment. Porto Alegre reached long-term correlation peak during the Games for all Twitter metrics (typically with $H > .85$). Its lowest levels occurred during the post-Games period with two metrics (number of hashtags and number of words) approaching randomness. Belo Horizonte also reached long-term correlation peak during the Games, especially for tweets and users, and less sharply for retweets, hashtags, and chatters (which stayed fairly flat). Persistence in the retweet metric continued to rise post-Games even while two metrics (tweets and users) decreased sharply. The pattern for Rio de Janeiro was somewhat different from these. Only hashtags and words peaked during the Games. The values for the other metrics were, in fact, at a minimum as the event was ongoing. No metrics approached randomness, with only words falling below $H = .8$. Rio's values were generally high overall. Despite these differences, the commonality of the level of persistence during the Games for all three cities is notable.

Figure 13 presents the DFA for 9 equal 17-day windows. The results largely echo those for the coarser partitioning. The levels of long-term correlation converge in all the cities for all the Twitter

metrics during the 17 days of the Games (which is the same period as described above). However, the path that leads to this point is, with few exceptions, non-static in all cities and characteristically different for Rio de Janeiro. For Rio de Janeiro, which has higher values for tweets, retweets, chatters, and active users, the Games and its adjacent windows reveal decreased long-term correlation in the patterns of Twitter activity for all metrics but the average of hashtags and words per post. Long-term correlation is greatest further away from the Games, before and after the event. At the end of the scraping period, Rio seems to have increased its patterns of long-term correlation, returning to a level previously observed. The Games seem to have the opposite effect for Belo Horizonte and Porto Alegre: Long-term correlation increases sharply in all metrics during the Games in comparison to the previous window. Belo Horizonte also seems to stand out with respect to the average number of hashtags and words per post, using more of each than the other two Brazilian cities.

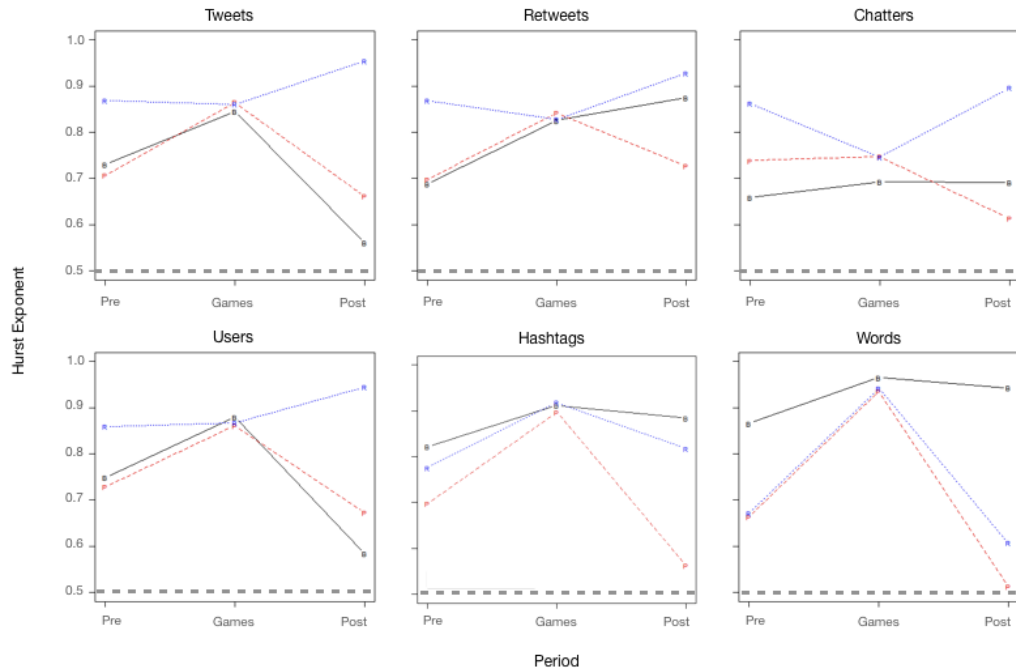


Figure 12. Hurst exponents derived from DFA for (a) tweets, (b) retweets, (c) chatters, (d) users, (e) average number of hashtags, and (f) average number of words pre-, during, and post-Games for Rio de Janeiro (blue dotted line with R markers), Belo Horizonte (black line with B markers), Porto Alegre (red dashed line with P markers). Dashed grey line indicates the threshold for randomized long-term correlation.

The first 17 days after the Games are mostly followed by a decrease in the long-term correlation in all cities, and in all cases, with some Twitter metrics approaching randomness more than others. The H levels achieved during the Games are maintained only for Rio de Janeiro's number of tweets, retweets, and users. The observed decrease is reversed typically after 17 and 34 days after the end of the Games for most of the Twitter metrics analyzed and the cities compared.

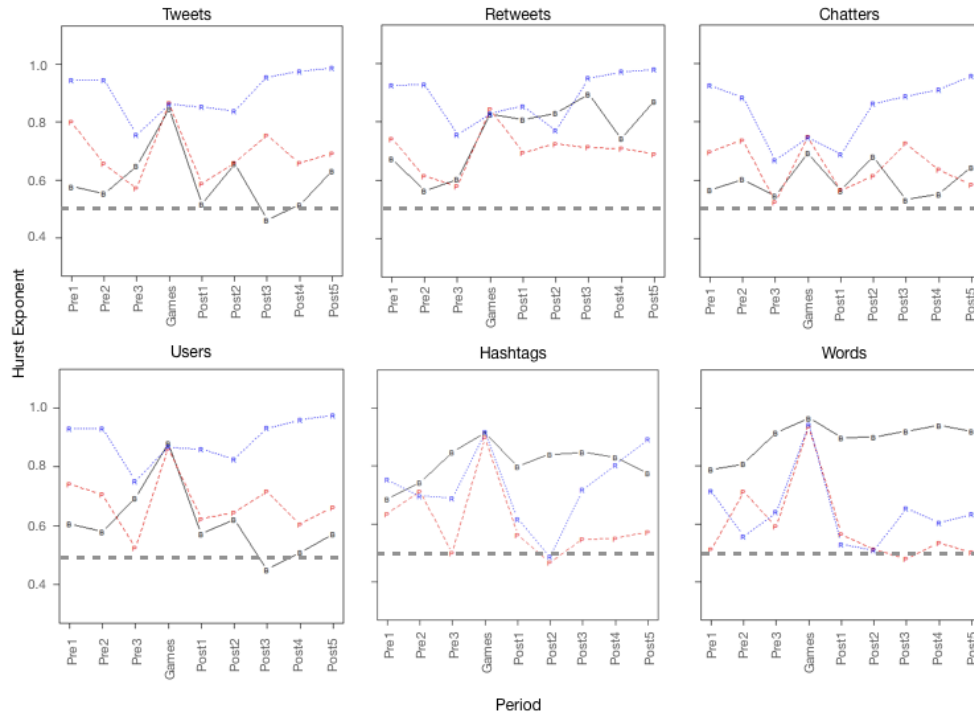


Figure 13. Plots of DFA for 17 days windows of tweets, retweets, chatters, users, average number of hashtags, and average number of words per tweet. X-axis in sequence: Three pre-Games windows (starting 51, 34, and 17 days before games), Olympic Games window, five post-Games windows (ending 17, 34, 51, 68, and 85 days after games). Cities: Rio de Janeiro (blue dotted line with R markers), Belo Horizonte (black line with B markers), Porto Alegre (red dashed line with P markers). Dashed grey line indicates the threshold for randomized long-term correlation.

In short, DFA does not indicate a legacy of the Games in the form of a change in long-term correlation. Regardless of a city's functional distance from the Games, post-Games states are very close to pre-Games states. It may, indeed, be the case that the event did not affect the dynamics of people's posts about LTPA.

FFT

The frequency analysis of Belo Horizonte and Porto Alegre echoes that for Rio de Janeiro, yielding the 22.5 or 25.7-hr periods in all windows (Figure 14), approximating the 24-hr daily pattern expected for Twitter data. The 12-hr period is also above threshold in almost all windows, with a handful of exceptions: windows 1 and 3 for Porto Alegre, and windows 1 and 5 for Belo Horizonte were below threshold; window 3 for Belo Horizonte yielded a slightly higher period of 13.8 hr.

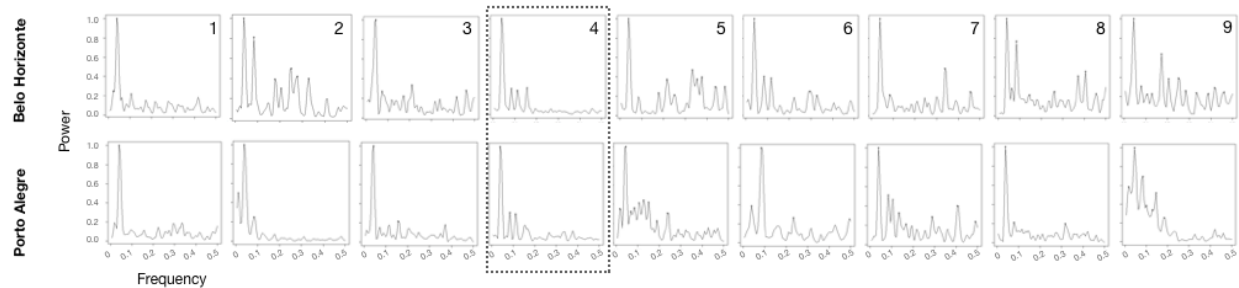


Figure 14. Periodograms generated from the FFT for 17-day windows of tweet activity in Belo Horizonte and Porto Alegre. (1-3) Pre-Games windows starting at 51, 34, and 17 days before the Games, (4) Olympic Games window, (5-9) post-Games windows ending at 17, 34, 51, 68, and 85 days after games).

Periods substantially higher than 24-hr were observed in a few windows: for Belo Horizonte, 45-hr was apparent in window 1; for Porto Alegre, 60-hr was apparent in windows 5 and 9, and 90-hr in window 2; and for Rio, 90-hr was apparent in window 5.

Above-threshold high frequencies were also apparent: in Belo Horizonte windows 2 and 3, and 5-9, and 5-8 in Porto Alegre. This may indicate that people are engaging more often in tweeting about LTPA. The periods 9 and 6.4 hr apparent in window 4 (i.e., during the Games) for Rio, also show up for Belo Horizonte, as does the 9-hr period for Porto Alegre. These may indicate that tweeting inside the network tends to follow more the work/rest/recreation cycle.

In summary, whereas the signal of tweet activity for Rio is composed of only a few important frequencies (up to 4), the final signals for Belo Horizonte and Porto Alegre have many more frequencies

(up to 12). This may indicate a tendency towards more random behavior by these systems with respect to the topic of LTPA.

Wavelet and Cross-wavelet Coherence

The Wavelet Power Analysis performed on the entire time series for the two Brazilian cities reveals a similar pattern to the one seen for Rio de Janeiro (Figure 15). During the Games the high-power levels (significantly different from white noise) across the spectrum are discernible indicating that many frequencies are contributing to the signal. Although the main frequencies of 24 and 12 hr are still apparent, they are only significant during the Games. Once again, the finer detail available in the sequence of 17-day windows shows high power levels around the 24 and 12-hr periods, but not with the same predominance as seen in Rio de Janeiro (Figure 16). It is not possible to visually identify the impact of the Games in the LTPA communication in these two periods for windows 3, 4 and 5. What is apparent is that the last 100 hr before the event in Belo Horizonte, and the first 100 hr and last 200 hr during the Games in both cities generate high power, and significantly so, through many periods of the same time frame. The same occurrence of multiple simultaneous significant high-power level frequencies can also be identified in windows 1 and 9 in Belo Horizonte, and 2, 5, and 9 in Porto Alegre.

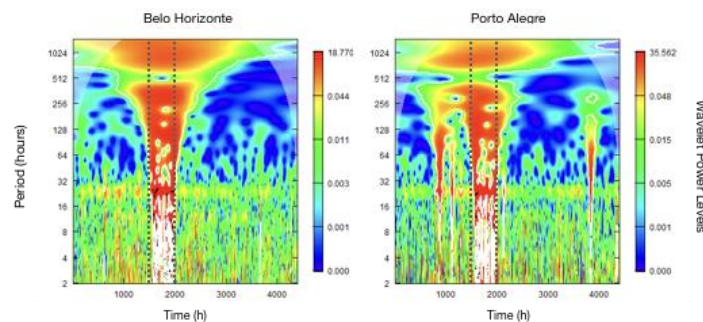


Figure 15. The Period \times Time Wavelet Power Spectrum generated for the entire period of tweet activity collected in Belo Horizonte and Porto Alegre (a total of 4391 hr). The Y-axis shows the candidate frequencies; the X-axis shows the time from the beginning to the end of the time series; color indicates the power level. Vertical dashed lines delimit the occurrence of the Games (between 1560 and 1967 hr).

Beyond the 24-hr period, no other high period has enough magnitude to be considered a contributor to the final time series of tweet production. On the other hand, the fast frequencies change from window to window, showing persistence across periods (from 1 to 24-hr) and fewer high-power significant occurrences over time. Especially in windows 3 and 4, the occurrences of high magnitude fast frequencies are much more concentrated at certain times, meaning that many frequencies have high power and contribute to the final signal.

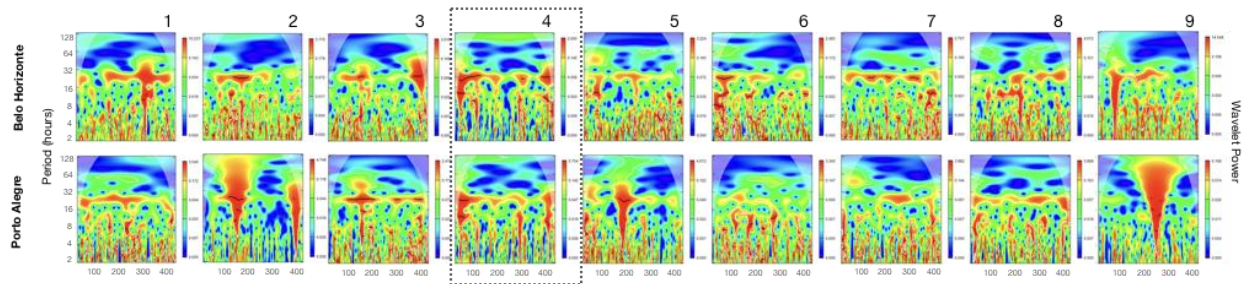


Figure 16. Wavelet Power Spectrums generated for 17-day windows of tweet activity in Belo Horizonte and Porto Alegre. The Y-axis shows the candidate frequencies; the X-axis shows the time from the beginning to the end of each window. (1-3) Pre-Games windows starting at 51, 34, and 17 days before the Games; (4) the Olympic Games window; (5-9) post-Games windows ending at 17, 34, 51, 68, and 85 days after the Games. (Regions that are significantly different from white noise are outlined in white.)

A frequency coordination analysis of the tweet activity in terms of the cross-wavelet coherence between Rio de Janeiro and the targeted cities confirmed the insights allowed by the wavelet analysis. This can be seen in Figure 17 and in Table 2. In the plots, significant moments of mostly in-phase entrainment of the cities' routines can be seen during the Games and its boundaries throughout all the periods, identified by the red colors delimited by a black line. Although Belo Horizonte's coordination with Rio de Janeiro is slightly prolonged compared to that of Porto Alegre, the coherence of both cities is quite similar. Periods in the proximity of 24 hr show strong coherence, mostly in phase, across the entire time series. The strength of the correlation during the Games and its surroundings is remarkable. The 12-hr period was also present, but with less coherent moments. Around the 256-hr period, strong correlations that go beyond the occurrence of the Games are also apparent, with greater prominence pre-Games in Belo Horizonte.

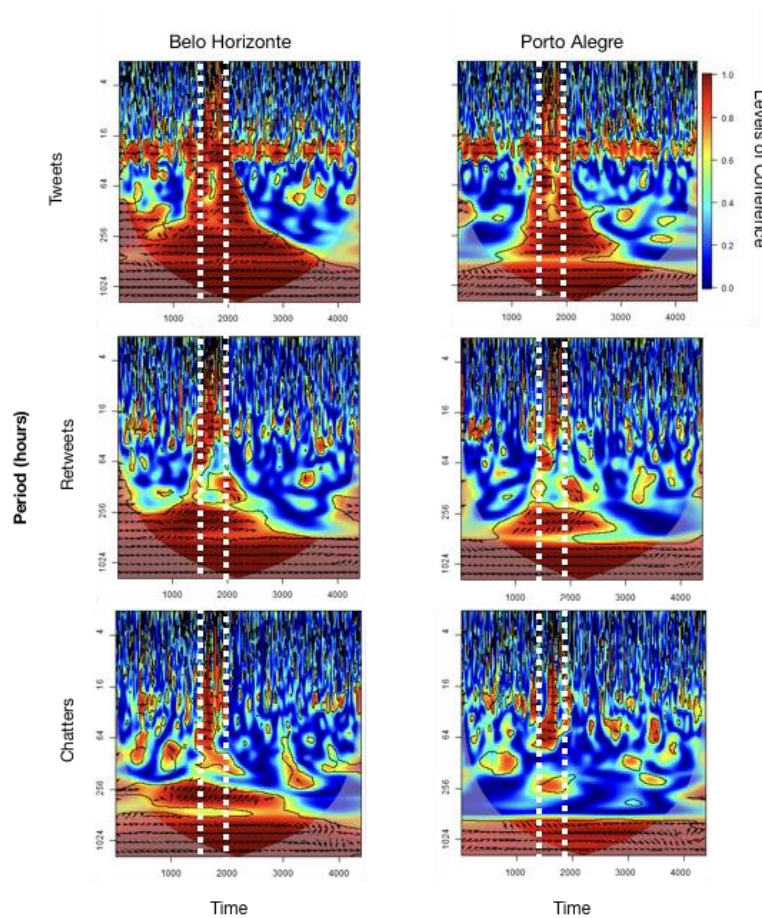


Figure 17. Cross-wavelet coherence of activity in Rio de Janeiro and (left) Belo Horizonte and (right) Porto Alegre during the entire scraping period for (top) tweets, (middle) retweets, and (bottom) chatters. The Y-axis shows the candidate frequencies; the X-axis shows the time from the beginning to the end of the time series (a total of 4391 hr). Vertical dashed white lines delimit the occurrence of the Games (between 1560 and 1967 hr).

The Cross-wavelet Coherence Analysis was further investigated for the 6, 12, and 24-hr periods (Figure 18 and 19) during three 17-day windows: pre- (from 51 to 34 days before the Games), during, and post- (from 34 days to 51 days after the Games). The results presented in Table 2 complement the cross-wavelet coherence plots. The pre- and post- windows are equally far from the Games and allow examination of how the Games changed communication about LTPA. The percentage of occurrence of significant coherence between Rio de Janeiro and its Brazilian counterparts is high at all times for the period of 24 hr. During the Games this coherence reaches 100% significance in both cities. This is quite

apparent in the cross-wavelet coherence plots in the predominance of red areas encircled by black lines (Figure 17, 18-C, and 19-C).

Table 2
Cross-Wavelet Coherence of Tweets (with Relative Phase) between Rio De Janeiro and Its Brazilian Counterparts as a Function of Epoch Relative to the Olympic Games for Each of Three Periods.

City	Epoch	Period (hr)	% Significance	Relative Phase M (SD)
Belo Horizonte	51-34 days prior	6	20.10	-0.24 (1.81)
		12	27.02	0.36 (1.32)
		24	87.50	0.02 (0.34)
	17 days during	6	85.05	0.10 (0.71)
		12	88.24	-0.02 (0.31)
		24	100.00	0.15 (0.19)
	34-51 days after	6	4.41	0.14 (1.73)
		12	15.44	-0.91 (1.24)
		24	76.00	0.14 (0.58)
Porto Alegre	51-34 days prior	6	19.12	0.18 (1.89)
		12	27.70	-0.17 (1.18)
		24	94.40	0.12 (0.16)
	17 days during	6	68.14	0.17 (0.82)
		12	83.11	-0.02 (0.55)
		24	100.00	0.07 (0.29)
	34-51 days after	6	15.70	-0.01 (1.58)
		12	28.70	-0.09 (1.15)
		24	96.81	-0.09 (0.35)

With respect to the 12-hr frequency, the percentage of significant coherence is above 83% during the Games for both Belo Horizonte and Porto Alegre and below 30% before and after the Games. Indeed, for Belo Horizonte, significantly coherent behavior declines by more than 11% from the pre- to post-Games windows. Porto Alegre, in contrast, did not show the same decline.

It is apparent that inphase is the common pattern of coordination with Rio de Janeiro. The last column of Table 2 indicates this in its phase average values, which are predominantly close to 0 over all analyzed epochs relative to the Games. Additionally, the occurrence of peaks around phase 0 in both cities' histograms points to the same conclusion—there is phase locking in these data (Figure 18-A and Figure 19-A). The predominance of the inphase coordination is found in the 12 hr-period in the pre- and

during Games windows for Belo Horizonte and during and post-Games windows for Porto Alegre. Moreover, the phase coordination in the 24-hr period does not occupy any other pattern far away from inphase. These concentrations are evidence of phase locking between the Brazilian cities and Rio de Janeiro. For the 6-hr period, the predominance of inphase coordination only arises during the Games; the pre- and post-Games patterns of coordination have no particular tendency in either city (Figures 18-B and 19-B). Although, the Games do not shift the relative phases (Figures 18 and 19, sections B), they increase the significant occurrence of coherent patterns among the cities (Sections C and D of Figures 18 and 19, and Table 2).

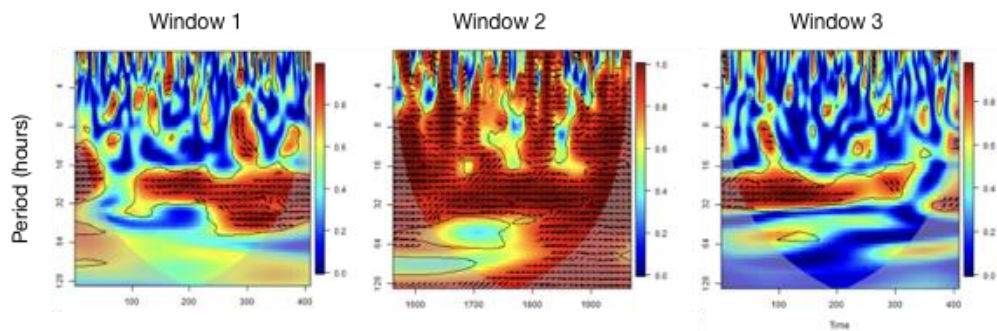


Figure 18. Cross-wavelet coherence of tweet production between Rio de Janeiro and Belo Horizonte with phase coordination for three windows of 17 days: (1) 51-34 days before the Games, (2) Games, (3) 34-51 days after the Games.

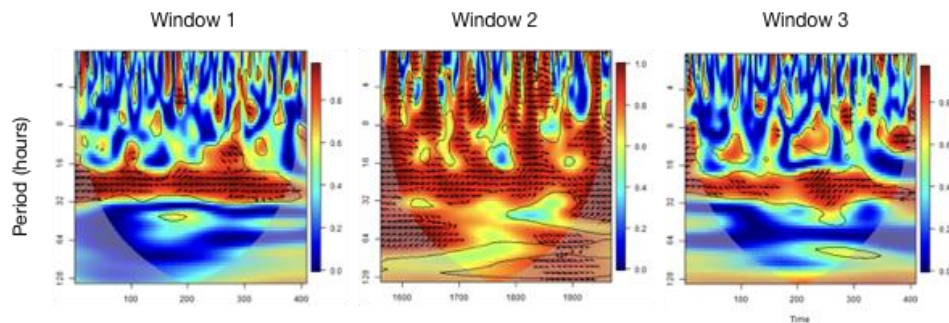


Figure 19. Cross-wavelet coherence of tweet production between Rio de Janeiro and Porto Alegre with phase coordination for three windows of 17 days: (1) 51-34 days before the Games, (2) Games, (3) 34-51 days after the Games.

Summary

Clearly an influence of the Games on LTPA Twitter activity was felt in all of the Brazilian cities. Not surprisingly, the most dramatic effect was apparent in the principal host city. But the other Brazilian

cities show a similar kind of contagious growth in the cumulative sum of many Twitter metrics. And all three cities converged on the same degree of persistence during the Games. Although Belo Horizonte and Porto Alegre generally have many more frequencies than Rio de Janeiro, they echoed the appearance of shorter frequencies and disappearance of longer frequencies during the Games. And the power of the frequencies was elevated during the Games. Finally, the coherence of activity with Rio de Janeiro increased dramatically during the Games. The evidence for a legacy is minimal, being limited to a small increase in coherence in the DFA for Rio de Janeiro after the Games.

CHAPTER 5

Study 2: Social Legacy vs. Time Zone Proximity

If sports engagement is a legacy of a mega-sporting event, then one might expect it to be made visible through more frequent LTPA tweets in former host cities relative to non-host cities, particularly before the event starts. One reason is that hosting-related changes in those localities might be expected to create new possibilities for action—new affordances (Gibson, 1979; also, see General Discussion)—that might promote LTPA communication about those actions. The Games might also reinforce the tendency of sports engagement, resulting in an augmented frequency of tweets after the Games when compared with before the Games.

Study 2 examined the frequency of tweets during the Rio Olympics for two former host cities, Atlanta 1996 and Sydney 2000, and two non-host control cities from the same countries and time zones as the host cities, respectively, Philadelphia and Brisbane. Whereas Atlanta and Philadelphia are in essentially the same time zone as Rio (1 hour earlier), Sydney and Brisbane are 13 hours ahead. All things being equal, it might be expected that the closer the targeted city time zone to the host city, the easier it would be for Twitter activity during the Games to be accentuated as communication occurs at compatible times. A social legacy, however, would suggest that the hosting history would be important over and above time zone. Twitter users from former host cities might be “primed” by those experiences such that the current social network strengthens the objective necessity of LTPA for those users.

Descriptive Statistics

The analysis considered a total of 102,459 tweets in Atlanta; 33,244 in Philadelphia; 37,378 in Sydney; and 13,242 in Brisbane. In terms of absolute numbers, Atlanta tweets more than the other cities (Table3). Its rate is double that of Rio with less than half its population (considering each municipality

plus its metropolitan area). Total LTPA tweets during the scraping period as a percentage of population were 1.95 for Atlanta, 0.82 for Sydney, 0.59 for Philadelphia, and 0.59 for Brisbane. Again, to facilitate comparison to each other as well as to Rio de Janeiro, these percentages reinforce the need to normalize each city's data relative to its own Twitter habit.

Table 3
Mean Number of Tweets per Hour per Day (with SD) during Three Epochs Relative to the Olympic Games for Four Cities as a Function of Time Zone (Relative to Rio de Janeiro) and Former Host Status.

City	Time Zone	Host Status	Epoch	Tweets
Atlanta	-1 hr	Former	51-34 days prior	22.32 (10.94)
			17 days during	26.94 (12.38)
			34-51 days after	25.16 (11.13)
Philadelphia	-1 hr	Never	51-34 days prior	7.41 (5.46)
			17 days during	9.26 (7.80)
			34-51 days after	7.56 (6.47)
Sydney	+13 hr	Former	51-34 days prior	7.95 (5.06)
			17 days during	12.80 (10.43)
			34-51 days after	8.25 (4.73)
Brisbane	+13 hr	Never	51-34 days prior	2.89 (2.40)
			17 days during	4.81 (4.77)
			34-51 days after	2.59 (2.31)

As in Study 1, the concentration of the tweet activity of each city in Figure 20 and Table 3 coincides with the epoch of the Olympic Games. However, this increase is not as dramatic as it was for the Brazilian cities, and is characterized by very different proportions of tweets, retweets, and chatters. Figure 21 shows the proportional daily Twitter activity of tweets, retweets, and chatters produced throughout the scraping period. As in the Brazilian cities of Study 1, individual Tweets comprise the bulk of activity. However, the contribution of retweets to the network is not as straightforward. Atlanta and Sydney seem to share the familiar pattern of an increase in the proportion of retweets during the Games, followed by a subsequent decrease to pre-Games levels after the event. This pattern is less clear for Philadelphia. The pattern for Brisbane is quite distinct: It is characterized by a much lower proportion of tweets overall than the other cities, with a correspondingly high proportion of retweets. The latter decreases to the level of the proportion of chatters in mid-September, accompanied by a slight increase in the proportion of tweets.

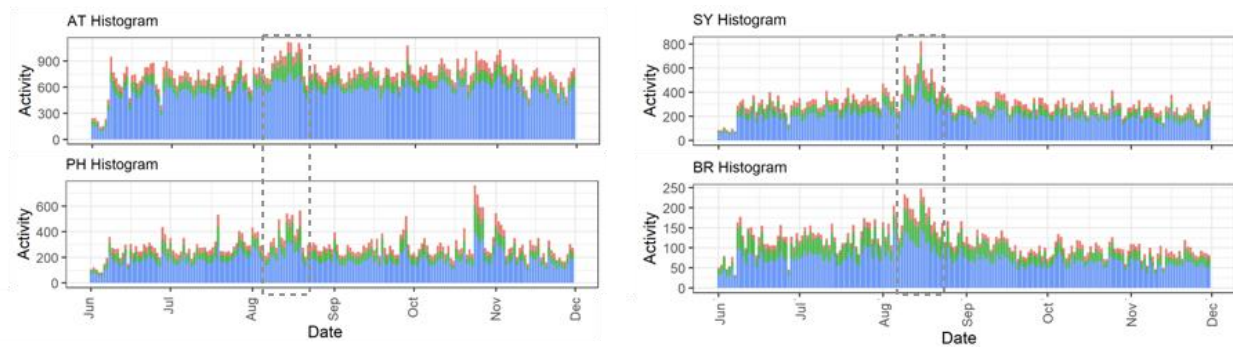


Figure 20. Histograms of daily LTPA-related tweets (blue), retweets (green), and chatters (pink) over the entire period of data collection for Atlanta (AT), Sydney (SY), Philadelphia (PH), Brisbane (BR).

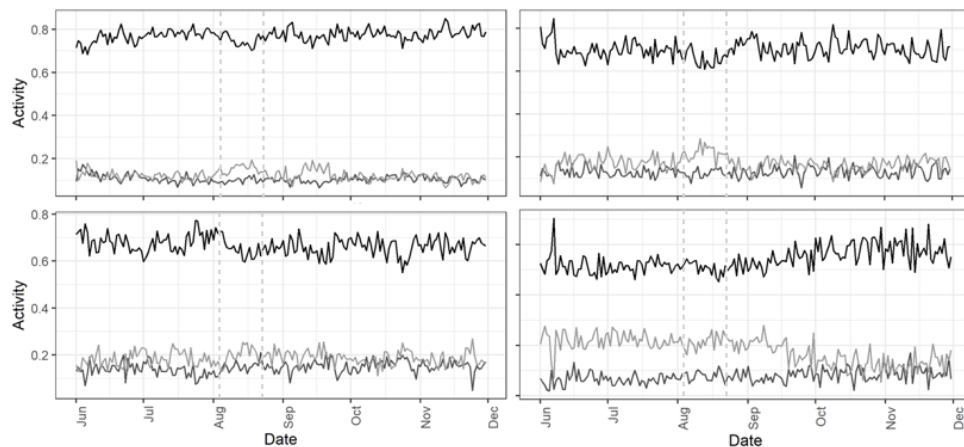


Figure 21. Time series of proportional daily LTPA-related Twitter activity metrics in Atlanta (top-left), Sydney (top-right), Philadelphia (bottom-left), and Brisbane (bottom-right): Number of tweets (black), retweets (light grey), and chatters (medium grey). Vertical dashed lines delimit the occurrence of the Games.

The weekly patterns of tweet production (Figure 22) reveal variability beyond ± 1 SD during the Games in all targeted cities. This behavior remains for some weekdays in Atlanta and Philadelphia after the Games. Weekend activity is constrained around the mean with less variability in all cities except for pre-Games in Brisbane. Unlike Rio but like the other Brazilian cities, there is great variability from the mean beyond 1 SD during the Games. With respect to the three patterns of normalized average weekly Twitter activity identified in Morales et al. (2017), the North American cities should resemble the graduated double-peak of the Brazilian cities while the Australian cities might be expected to display a pattern with a single large peak of activity during the day that is typical of Asian and Oceanian cities. However, this does not seem to characterize the patterns for the LTPA topic. If anything, a double-peak

pattern is observed on some weekdays before and after the Games for the Australian cities whereas the North American cities show more of a single peak.

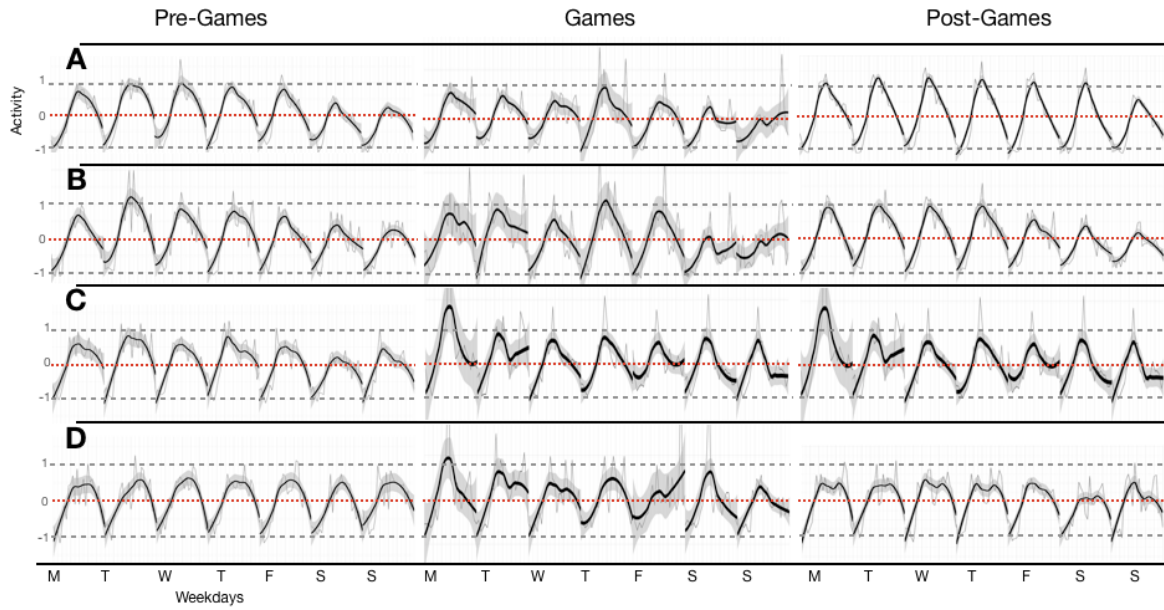


Figure 22. Weekly LTPA-related tweets for Atlanta (A), Philadelphia (B), Sydney (C), and Brisbane (D). Pre-Games (left column), Games (middle column), and (right column) post-Games. Activity is averaged over corresponding hours during each day of the week throughout the corresponding scraping period. Data were normalized (z-scores) for the entire scraping period. Red dashed lines indicate the mean; the grey dashed lines indicate ± 1 SD.

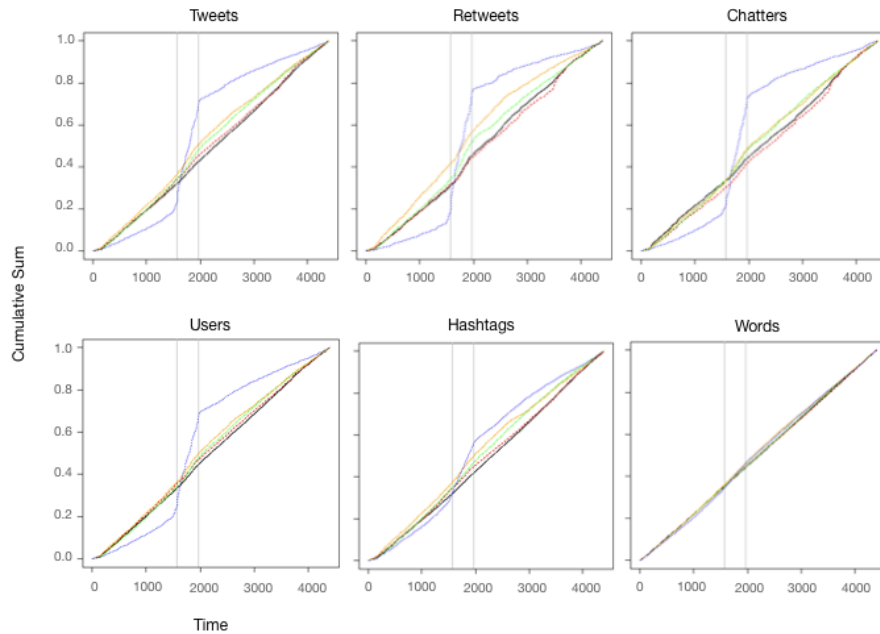


Figure 23. Cumulative Sums for tweets, retweets, chatters, users, average number of hashtags, and average number of words during the entire period of data collection for Rio de Janeiro (blue dotted line), Atlanta (black line), Philadelphia (red dashed line), Sydney (green dot-dashed line), and Brisbane (orange two dashed line). Data are normalized by each city's total number of tweets during the scraping period. The occurrence of the Games is bracketed by the dashed vertical lines.

The cumulative sums of the Twitter metrics are shown in Figure 23, again scaled by the total activity during the scraping period. Relative to the baseline provided by Rio de Janeiro, no other city—regardless of hosting history or time zone—presents such a dramatic exponential alteration of behavior. All cities (including Rio) show linear growth in the number of words. Their cumulative sum slopes are very similar, but especially for tweets, retweets, and hashtags, it is possible to identify that (i) Sydney and Brisbane, (ii) Atlanta and Philadelphia, and (iii) Rio de Janeiro have very specific slope characteristics. Visual inspection suggests that LTPA is not a topic that is impacted in a contagious fashion by the Games, except in the case of the host city. To test this assumption more formally, however, the rate of growth will be examined in a later section, *Comparisons of all cities*.

DFA

Figure 24 presents results of the DFA for the entire scraping period before the Games, the 17 days of the Games, and the entire scraping period after the Games. Values of H are between .5 and 1.0 for all metrics throughout the scraping period, indicating persistence. And although the degree of persistence in the three cities varied during the pre- and post-Games periods, the values of H converged for tweets, chatters, users, and average number of words per tweet during the Olympic Games, although not as tightly as for the Brazilian cities.

The pre-Games values for number of tweets and users do not change during the Games but they diverge after the Games. After the Games, long-term correlation for Rio de Janeiro increases even more, Atlanta and Sydney values are maintained, and Philadelphia and Brisbane experience a decrease, suggesting that previous hosting status, regardless of time zone, is an important factor in this matter. In the case of retweets and chatters, the North American cities seem to follow a similar pattern distinct from the Australian cities, suggesting the importance of a contiguous time zone in the more social aspects of Twitter. Rio de Janeiro's drastic decrease in long-term correlation for average number of

words after Games is not echoed by any other targeted location. The impact of being the current host seems to constrain the number of words written in posts.

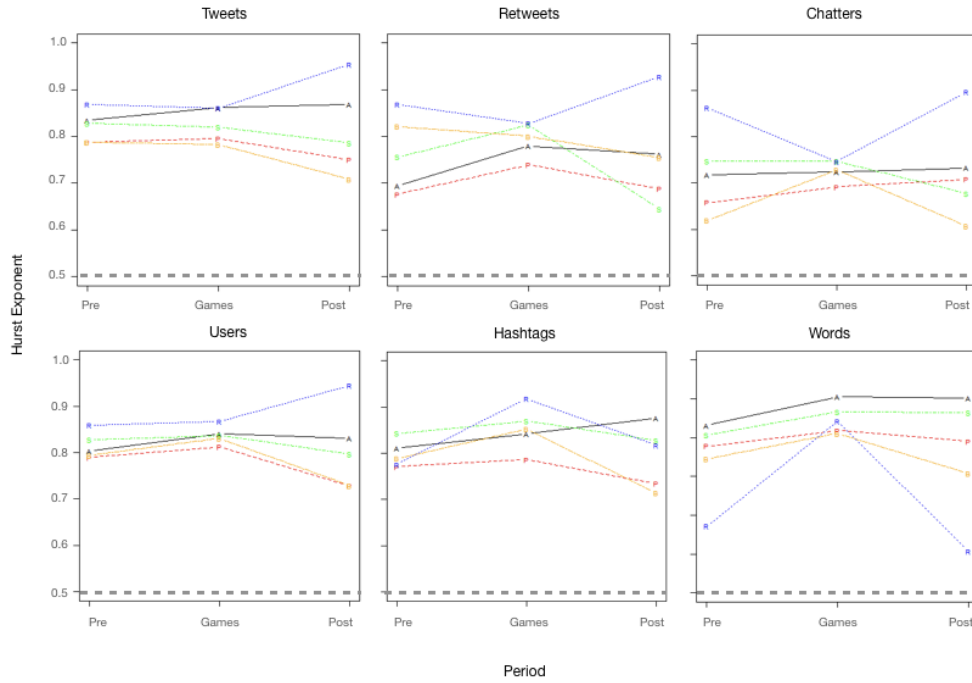


Figure 24. DFA plots for tweets, retweets, chatters, users, average number of hashtags, and average number of words activities, before, during, and after the Games. Cities: Rio de Janeiro (blue dotted line), Atlanta (black line), Philadelphia (red dashed line), Sydney (green dot-dashed line), Brisbane (orange two dashed line). Dashed grey line indicates the threshold for randomized long-term correlation.

Figure 25 presents the DFA for 9 equal 17-day windows. The analysis shows that the lower long-term correlations observed in Rio de Janeiro during the Games and in the windows surrounding the Games for tweets, retweets, chatters, and users, is not seen in any other city. Instead, long-term correlation increases for all the targeted cities during the Games, with the increase building up from some windows pre-Games, and sustaining itself sometimes for longer, after the end of the event. Philadelphia's behavior matches Atlanta's for chatters and users, and Brisbane's behavior matches Sydney's for tweets, retweets, users, and hashtags. Rio de Janeiro also shows a distinct pattern in its increase in long-term correlation for the average number of hashtags and words. While other cities show a build up in dependency evolving from previous windows, none of them is as drastic and exponential as the pattern observed in the host city. Philadelphia echoes the drastic drop in long-term correlation in the post-Games windows, however, in the average hashtags metric. Moreover, for both

Rio de Janeiro and Philadelphia, the precipitous drop is followed by an increase built up sequentially through the post-Games windows. For Sydney and Brisbane there is a relaxation stage in the post-Games windows with a decrease of dependency in a sequential manner. Atlanta's dependency levels are fairly flat for the number of tweets, users, average number of hashtags, and words, and Sydney's is flat for the average number of words. In these cases, it seems that the Games do not disrupt the dynamics of long-term correlations. It is also interesting to note, however, that for such a stable dynamics locality as Atlanta, the Games seem to have influenced the reproduction of information through retweets, and of addressing others in the conversation through chatters.

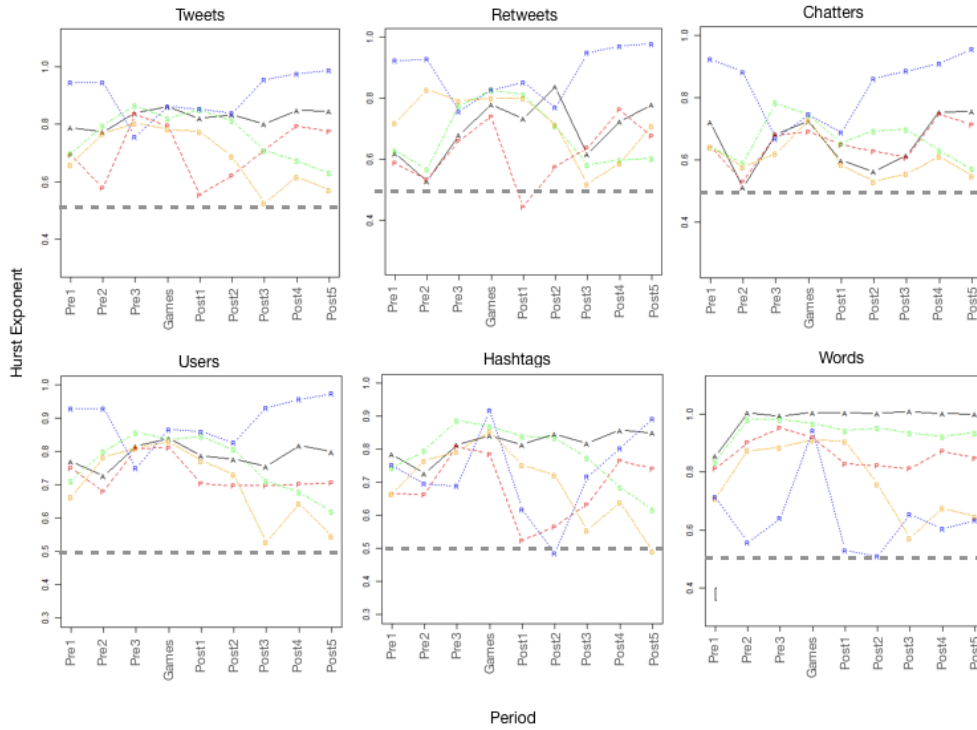


Figure 25. DFA plots for 17-day windows of tweets, retweets, chatters, users, average number of hashtags, and average number of words per tweet. X-axis in sequence: Three pre-Games windows (starting 51, 34, and 17 days before the Games), the Olympic Games window, five post-Games windows (ending 17, 34, 51, 68, and 85 days after the Games). Cities: Rio de Janeiro (blue dotted line), Atlanta (black line), Philadelphia (red dashed line), Sydney (green dot-dashed line), Brisbane (orange two dashed line). Dashed grey line indicates the threshold for randomized long-term correlation.

FFT

The frequency analysis again yields the periods of 22.5 or 25.7 hr in all windows, approximating the 24-hr daily pattern (Figure 26) expected for Twitter data. The 12-hr period that was clear in Rio was present in all of Sydney's windows, and in most of the windows of the other cities.

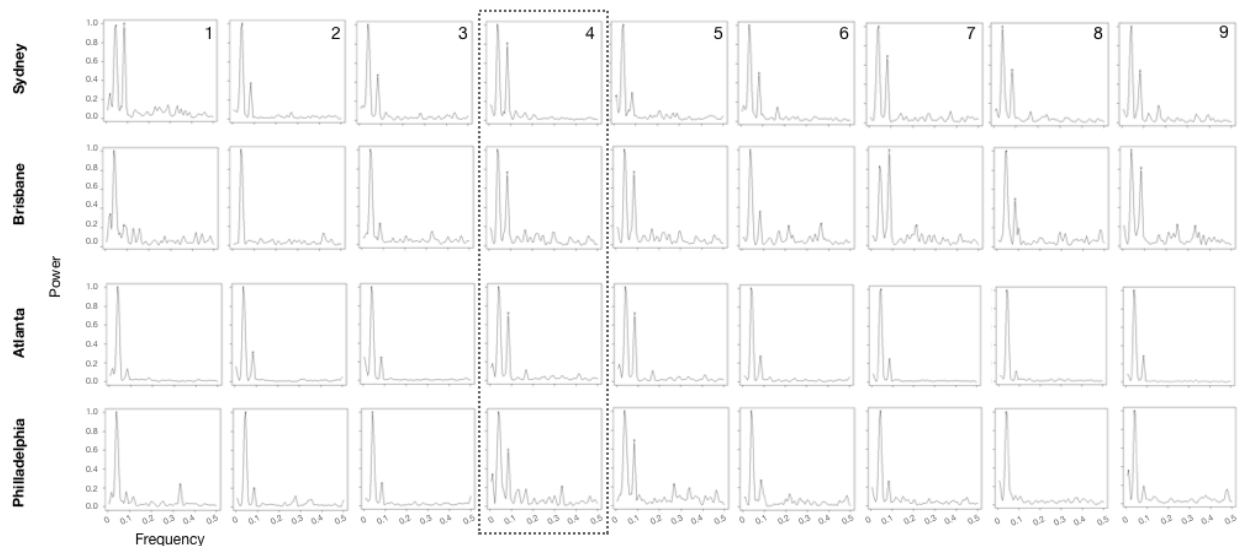


Figure 26. Periodograms generated from the FFT for 17-day windows of tweet activity in Sydney, Brisbane, Atlanta, and Philadelphia. (1-3) Pre-Games windows starting at 51, 34, and 17 days before the Games, (4) Olympic Games window, (5-9) post-Games windows ending at 17, 34, 51, 68, and 85 days after games.

Focusing on the Australian cities, Sydney produced the least power for the 12-hr period localized at the boundaries of the Games (windows 2, 3, and 5) while the greatest power for that frequency occurred in windows 1, 4 and 7. For Brisbane, while there was no 12-hr period in window 2 (34-18 days before the Games), the power of the 12-hr period was considerable during and after the Games (windows 4, 5, 7, and 9). Sydney sometimes showed slower frequencies (60 and 90 hr in windows 1 and 5), but no frequencies faster than 12 and 24 hr. Brisbane yielded a mix: a slower frequency of 45 hr in window 1; faster frequencies of around 4.5 hr identified in windows 5, 7, and 9; and the even faster frequency of 2.7 hr right after the Games.

Among the United States cities, the 12-hr period was only below threshold in windows 1 and 8 for Atlanta and in windows 1, 8, and 9 (i.e., far from the Games) for Philadelphia. Both cities showed the greatest power for the 12-hr period in windows 4 and 5. Philadelphia also showed a faster frequency (3 hr) in windows 1, 4, and 5, as well as a slower frequency (90 hr) in windows 4 and 9. Atlanta, in contrast, did not have frequency diversity apart from 24 and 12 hr.

While above-threshold high frequencies for Brisbane and Philadelphia could indicate that people are engaging more often in tweeting about LTPA during the Games, it could simply be a result of low tweet activity, which would promote the importance of a less powerful frequency. Differently from the Games period, LTPA communication in the investigated localities did not seem to follow the work, rest, and recreation cycle periods as seen in Rio, for example. It seems that the signal of tweet activity is composed primarily from two important frequencies. Brisbane and Philadelphia have more frequencies generating the final signal in some of the analyzed windows (up to 4).

Wavelet and Cross-wavelet Coherence

The Wavelet Power Analysis performed on the entire time series for the four targeted cities reveals patterns that are very different from Rio de Janeiro and the other Brazilian cities (Figure 27). The main 24-hr frequency is not only apparent in all cities, as in Study 1; it persists at a significant level throughout the scraping period with high power. Quite unlike the Brazilian cities, the epoch during the Games does not stand out in the spectrum. In both the U.S. and Australian cities, the Games can be detected in the concentration of significant white lines before 2000 hr.

Turning to the finer detail available in the Wavelet Power Analysis performed on the sequence of 17-day windows, the high-power levels around the 24-hr period are sustained more than the Brazilian cities, even Rio de Janeiro (Figure 28). The 12-hr period is less recurrent over time or across windows and localities, but still reaches power and significance. The impact of the Games is not as easily seen in

the spectrum as in Study 1. However, an increase in the occurrence of high-power fast frequencies across multiple periods (from 1–24 hr) is apparent in window 4 of all cities. In brief, at these moments, those fast frequencies become important and play a role in composing the final signal of the time series, or in other words, in generating the tweets. Also, the 12-hr period increases its power in window 4. In Atlanta and Philadelphia, it is apparent that the 24-hr period loses a little bit of its power and predominance after Games (window 5). No period higher than 24 hr has enough magnitude to be considered as contributing to the final time series of tweet production.

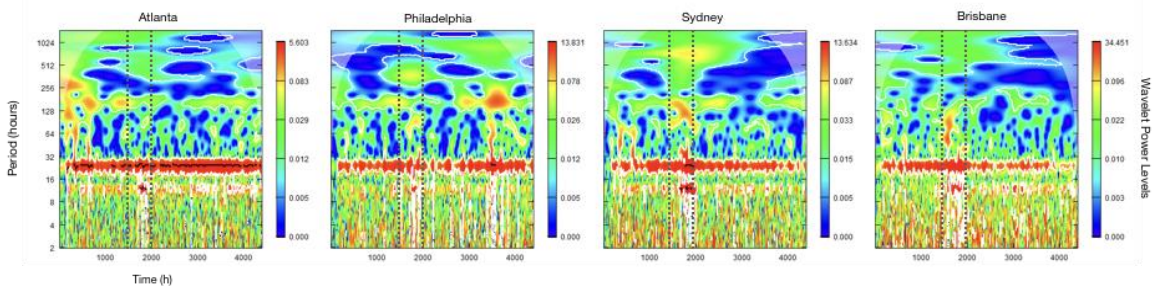


Figure 27. Period \times Time Wavelet Power Spectra generated for the entire period of tweet activity for Atlanta, Philadelphia, Sydney, and Brisbane (a total of 4391 hr). The Y-axis shows the candidate frequencies; the X-axis shows the time from the beginning to the end of the time series; color indicates the power level. Vertical dashed lines delimit the occurrence of the Games (between 1560 and 1967 hr).

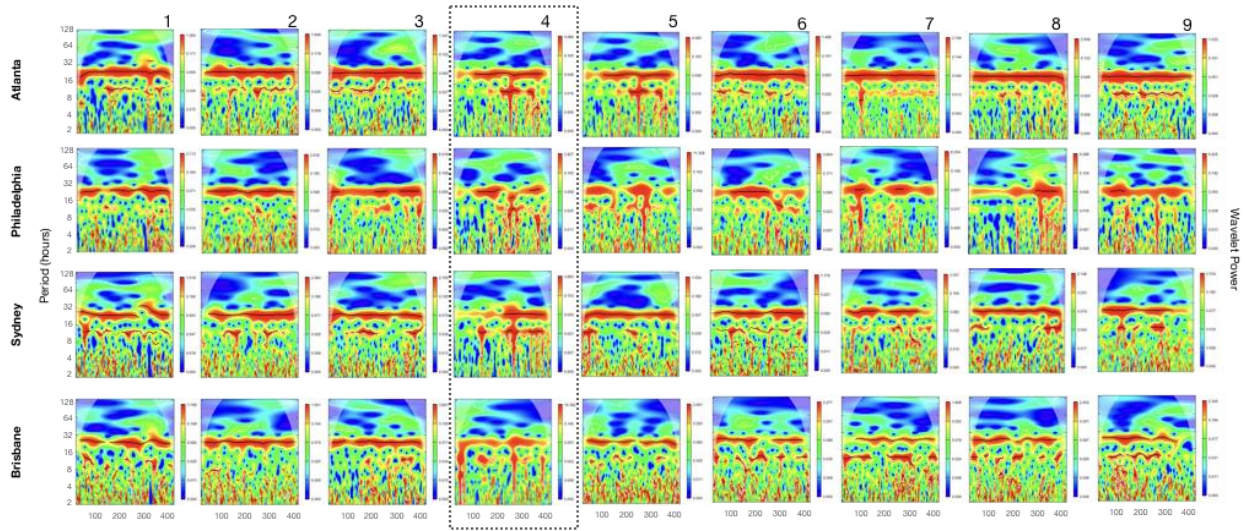


Figure 28. Period \times Time Wavelet Power Spectrums generated for successive 17-day windows of tweet activity in Atlanta, Philadelphia, Sydney, and Brisbane. (1-3) Pre-Games windows starting at 51, 34, and 17 days before the Games; (4) the Olympic Games window; (5-9) post-Games windows ending at 17, 34, 51, 68, and 85 days after the Games). The Y-axis shows the candidate frequencies; the X-axis shows the time from the beginning to the end of each window. (Regions that are significantly different from white noise are outlined in white.)

One goal of Study 2 was to assess whether being in the same time zone as the host city mattered to the effect of the Games on Twitter activity about LTPA. The Wavelet analysis suggests that it does not. The behavior of these 4 cities resembles each other, despite representing time zones that are 13 hr apart. Also, even though Rio de Janeiro and Atlanta and Philadelphia are in time zones that are only 1 hr apart, the U.S. cities do not respond in the same way to the Games as Rio de Janeiro or the other Brazilian cities analyzed.

The frequency coordination analysis in terms of the cross-wavelet coherence between Rio de Janeiro and the targeted cities confirmed the insights allowed by the wavelet analysis. This can be seen in Figure 29 and in Table 4. In the plots, the 24-hr period and its surroundings has a high significant correlation throughout the entire analyzed period. This effect was seen for number of tweets and chatters, with much more strength in the former case. All the cities presented a disruption in coherence of the 24-hr period close to the beginning and end of the Games, which could be an effect of the event. Note that inphase coordination here means, not that the cities' tweet production occurs at the same physical time, but in the same hour of the day, demarcating that the daily routines of the cities match Rio's.

The 12-hr period was also very prominent with respect to the number of tweets, providing evidence of coherence with Rio among all the targeted cities. This effect was not evident in the number of retweets or chatters. During the days of the Games, delimited by the white dashed lines, the extension of the 24-hr band is apparent, as is the emergence of a 256-hr coherent period in some of the cities.

Some inphase coherence is apparent before and after Games around the 256-hr period for all cities. Sydney and Philadelphia also show this phase coherence during the Games. In the case of Sydney, the phase coordination is shifted towards a more anti-phase relation, which might indicate that during the Games the tweets were being produced during the actual time of occurrence of the Games, which

reflects 13 hours of difference in the daily routine of the city. It might indicate that people changed their patterns of behavior during the Games, probably keeping up with the events and tweeting about it at non-conventional hours of the day.

Table 4
Cross-Wavelet Coherence of Tweets (with Relative Phase) between Rio De Janeiro and Four Target Cities as a Function of Time Zone (Relative to Rio de Janeiro), Former Host Status, and Epoch Relative to the Olympic Games for Each of Three Periods.

City	Time Zone	Host Status	Epoch	Period (hr)	% Significance	Relative Phase M (SD)
Atlanta	-1 hr	Former	51-34 days prior	6	28.20	0.35 (1.66)
				12	44.40	0.01 (1.08)
				24	100.00	-0.46 (0.19)
			17 days during	6	18.14	-0.29 (1.66)
				12	38.00	-0.14 (1.00)
				24	78.70	-0.98 (0.30)
			34-51 days after	6	15.20	-0.18 (1.94)
				12	56.62	0.25 (0.59)
				24	97.31	-0.55 (0.34)
Philadelphia	-1 hr	Never	51-34 days prior	6	11.30	0.06 (1.73)
				12	64.71	-0.21 (1.14)
				24	97.10	-0.54 (0.18)
			17 days during	6	18.90	-0.58 (1.62)
				12	45.10	-0.25 (0.96)
				24	63.00	-1.11 (0.49)
			34-51 days after	6	18.40	0.58 (1.83)
				12	32.60	-0.03 (0.93)
				24	85.30	-0.53 (0.36)
Sydney	+13 hr	Former	51-34 days prior	6	8.33	0.08 (1.68)
				12	54.17	-0.04 (0.93)
				24	52.21	-0.70 (0.72)
			17 days during	6	17.90	-0.07 (1.56)
				12	47.30	-0.22 (0.91)
				24	80.90	-1.54 (0.35)
			34-51 days after	6	11.03	0.48 (1.98)
				12	54.90	-0.95 (0.72)
				24	84.60	-0.52 (0.48)
Brisbane	+13 hr	Never	51-34 days prior	6	5.20	-0.02 (1.88)
				12	15.70	-0.46 (1.10)
				24	57.40	-0.45 (0.63)
			17 days during	6	11.52	-0.31 (1.77)
				12	62.80	-0.26 (0.84)
				24	87.80	-1.23 (0.36)
			34-51 days after	6	20.10	0.51 (1.88)
				12	73.00	-0.85 (0.83)
				24	73.80	-0.21 (0.29)

Both Australian cities show high significant coherence in the slower frequencies around the period of 1024 hr, which persists for a great portion of the time analyzed. This could be an indication of the

characteristic way these regions tweet but, due to the limitations of the analysis (notably, the cone of influence), not much information is available about the progression of coordination in this period.

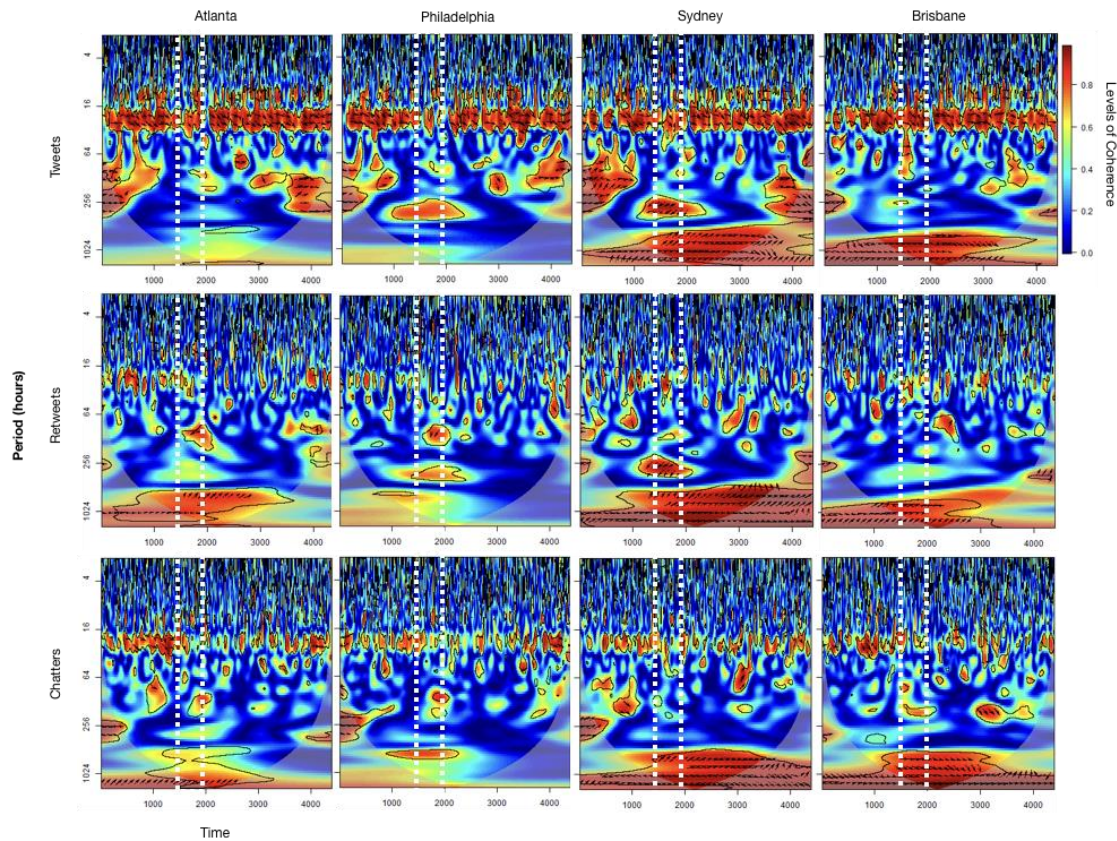


Figure 29. Cross-wavelet Coherence of activity in Rio de Janeiro with (left to right) Atlanta, Philadelphia, Sydney, and Brisbane during the entire scraping period for (top) tweets, (middle) retweets, and (bottom) chatters. The Y-axis shows the candidate frequencies; the X-axis shows the time from the beginning to the end of the timeseries (a total of 4391 hours). Vertical dashed white lines delimit the occurrence of the Games (between 1560 and 1967 hr).

The Cross-wavelet Coherence Analysis focused on the 6, 12, and 24-hr periods during three 17-day windows: pre- (from 51 to 34 days before the Games), during, and post- (from 34 days to 51 days after the Games). The results presented in Table 4 complement the cross-wavelet coherence plots.

Summary

Again, an influence of the Games on LTPA Twitter activity was felt in all of the targeted cities. However, it was not of the same magnitude as in the Brazilian cities. Contagious growth is not apparent

in the cumulative sums of the Twitter metrics (but, again, see the section *Comparisons of all cities*).

There is some convergence on the degree of persistence during the Games although not as sharp as for the Brazilian cities. But they seem to share neither the promotion of shorter frequencies nor the elevated power of frequencies during the Games that was seen in the Brazilian cities. Finally, systematicities as a function of time zone did not emerge.

The evidence for a legacy would be shown if Atlanta and Sydney were different from their non-host counterparts Philadelphia and Brisbane. But systematicities as a function of previous hosting status did not emerge.

CHAPTER 6

Study 3: Longevity of a Social Legacy

Hosting status did not provide evidence of a legacy in Study 2. As an alternative perspective on the issue, Study 3 was directed at whether the dynamic pattern of Twitter metrics was affected by how long ago a city hosted the Olympic Games. If sports engagement is a legacy of a mega-sporting event, perhaps that legacy gets reinvigorated in each Olympiad no matter where the Games were hosted each time. To address this, former host cities having native English or Portuguese languages were examined—London 2012, Sydney 2000, Atlanta 1996, Los Angeles 1984, and Melbourne 1956—and compared to Rio de Janeiro 2016.

A social legacy would be indicated by augmented frequency of tweets. As examples of possible effects, activity during the month immediately before the Games should exceed activity during the month prior to that; an overall increase should be apparent during the Games; or the level after the Games might be elevated relative to the pre-event level. In this case, the sports event could be considered a “Tweet attractor.” One might expect the strength of this attractor to decrease with the years (e.g., due to generational, infra-structure, economic and local political changes), in a relaxation fashion, meaning the time since hosting could matter to the frequency of LTPA tweets of the targeted cities.

Descriptive Statistics

The analysis considered 473,417 LTPA related tweets in London; 37,378 in Sydney; 102,459 in Atlanta; 501,413 in Los Angeles; and 41,141 in Melbourne. Table 5 shows the mean and standard deviation of tweets per targeted city per epoch. Total LTPA tweets during the scraping period as a percentage of population were 4.54 for London, .82 for Sydney, 1.95 for Atlanta, 4.07 for Los Angeles,

and .97 for Melbourne. Again, these percentages reinforce the need to normalize each city's data relative to its own Twitter habit.

Table 5
Mean Number of Tweets per Hour per Day (with SD) during Three Epochs Relative to the Olympic Games for Former Host Cities (with Host Year).

City	Epoch	Tweets
London (2012)	51-34 days prior	111.08 (55.89)
	17 days during	131.66 (56.92)
	34-51 days after	105.57 (44.48)
Sydney (2000)	51-34 days prior	7.95 (5.06)
	17 days during	12.80 (10.43)
	34-51 days after	8.25 (4.73)
Atlanta (1996)	51-34 days prior	22.32 (10.94)
	17 days during	26.94 (12.38)
	34-51 days after	25.16 (11.13)
Los Angeles (1984)	51-34 days prior	132.93 (44.85)
	17 days during	122.91 (23.94)
	34-51 days after	108.36 (22.05)
Melbourne (1956)	51-34 days prior	8.16 (8.23)
	17 days during	12.70 (11.71)
	34-51 days after	1.51 (8.96)

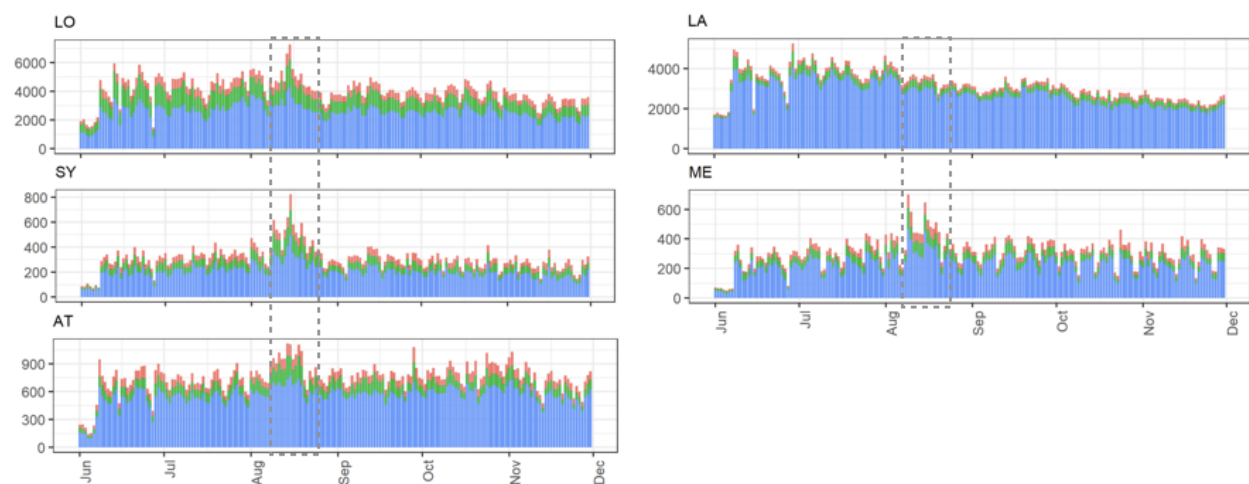


Figure 30. Histograms of daily LTPA-related tweets (blue), retweets (green), and chatters (pink) over the entire period of data collection for London 2012 (LO), Sydney 2000 (SY), Atlanta 1996 (AT), Los Angeles 1984 (LA), Melbourne 1956 (ME).

The pattern of daily LTPA-related tweets, retweets, and chatters is shown in Figure 30 and their proportions are shown in Figure 31. It is interesting to see that the disturbance seen in the pattern of Twitter metrics during the Olympic Games for the Brazilian cities is less clear in the former host cities. Los Angeles does not even show the characteristic increase in activity during the epoch of the Games, as

if the event did not interfere at all in people's routine of posts about LTPA. The increase in the proportion of retweets during the Games displayed by the Brazilian cities in Study 1 is very slight in Atlanta and Sydney but does not characterize Los Angeles, Melbourne, and London. London is the only city that presents non-overlapping lines for chatters and retweets throughout the scraping period (Figure 31). London also has the highest proportion of retweets and lowest proportion of tweets; Los Angeles is the opposite.

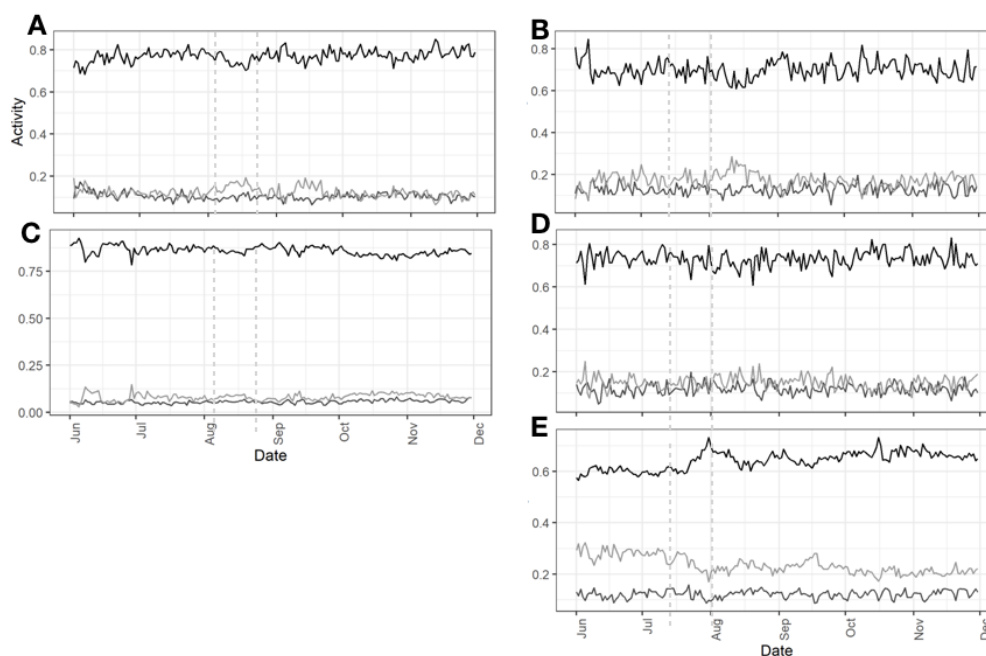


Figure 31. Time series of proportional daily LTPA-related Twitter activity metrics in Los Angeles (A), Sydney (B), Atlanta (C), Melbourne (D), and London (E): Number of tweets (black), retweets (light grey), and chatters (medium grey). Vertical dashed lines delimit the occurrence of the Games. For other targeted cities of Study 3 check plot in Study 2.

The weekly patterns of tweet production (Figure 32) reveal that, like all of the other targeted cities, the classic sleep time zones do not vanish during the Games as they did for Rio de Janeiro. For Los Angeles, tweet production is constrained in the period after the Games with a decrease in the hourly numbers of tweets from the mean shown by the extension of the curves towards -1 SD. During the Games a double-peak pattern, but not the expected graduated form, is revealed on Wednesdays,

Thursdays, and Saturdays. London pre-Games shows the equal double-peak pattern identified by Morales et al. 2017 which vanishes during the Games, reappearing in a very discrete manner in the after Games. As an Australian city, Melbourne might be expected to display a pattern with a prominent large peak of activity during the day that is typical of Asian and Oceanian cities. However, this does not seem to characterize the patterns for the LTPA topic. A pattern with less prominent double-peaks is apparent in the pre- and post-Games. The world characteristic pattern for Los Angeles is not evident in the LTPA weekly pattern. The weekends seem to have a common effect in all periods: activity is constrained around the mean.

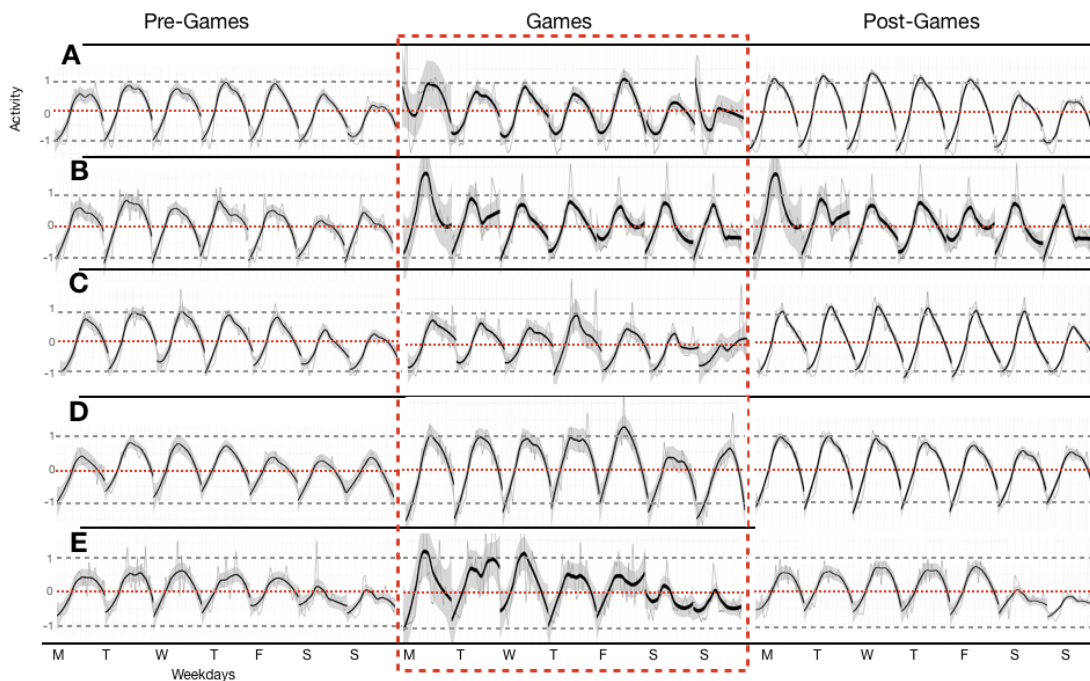


Figure 32. Weekly LTPA-related tweets for London 2012 (A), Sydney 2000 (B), Atlanta 1996 (C), Los Angeles 1984 (D), and Melbourne 1956 (E). Pre-Games (left), Games (middle), and (right) post-Games. Activity is averaged over corresponding hours during each day of the week throughout the corresponding scraping period. Data were normalized (z-scores) for the entire scraping period. Red dashed line indicates the mean; the grey dashed lines indicate ± 1 SD. For other targeted cities of Study 3 check plot in Study 2.

The cumulative sums of the Twitter metrics are shown in Figure 33, again scaled by the total activity during the scraping period. The curve for the contemporaneous host city is included for ease of comparison. The other cities present a mild alteration of behavior seen in the slope's steepness change during the Games. But no other city—regardless of recency of hosting—presents Rio's dramatic exponential alteration of behavior. All cities (including Rio) show linear growth in the number of words. Overall, Rio's slopes pre- and post-Games are also less steep than the other targeted cities. Again, visual inspection suggests that LTPA is not a topic that is impacted in a contagious fashion by the Games, except in the case of the host city. To test this assumption, however, an explicit examination of the rate of growth will be developed in the section that follows, *Comparisons of all cities*.

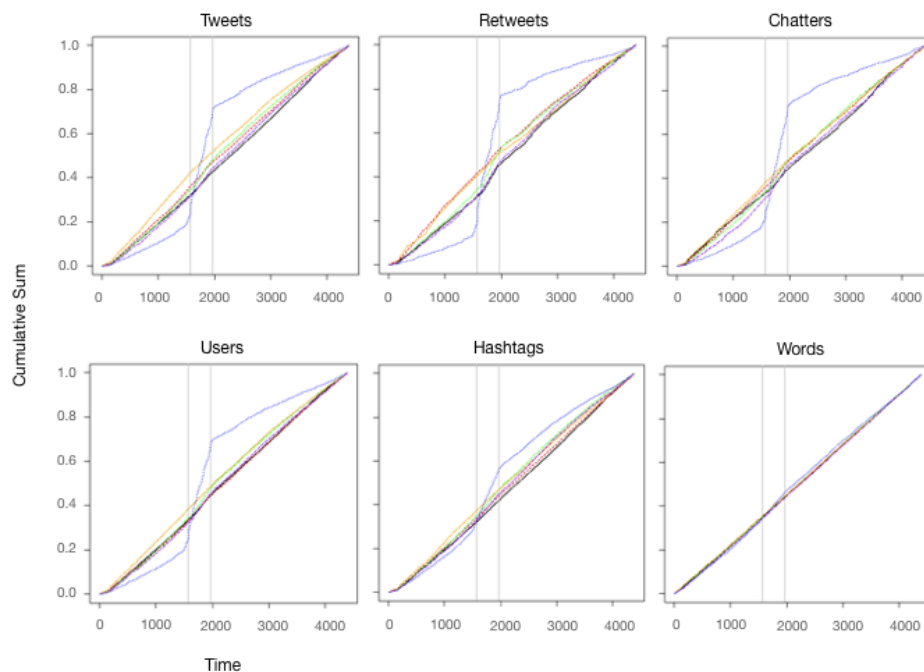


Figure 33. Cumulative Sums for tweets, retweets, chatters, users, average number of hashtags, and average number of words during the entire period of data collection for Rio de Janeiro (blue dotted line), Atlanta (black line), London (red dashed line), Sydney (green dot-dashed line), Los Angeles (orange two dashed line), Melbourne (purple long dashed line). Data are normalized by each city's total number of tweets during the scraping period. The occurrence of the Games is bracketed by the dashed vertical lines.

In terms of absolute numbers, London and Los Angeles are on a different scale from the other former host cities (see Appendix 2). Their descriptive statistics far exceed the others in almost all Twitter

related metrics. However, compared to Rio there are a few metrics during the Games that reach similar levels in these cities: maximum and standard deviation of tweets, retweets, chatters, and users.

Interestingly, Rio's average number of words per post sits well below all of the former hosts, only approaching their level during the Games (accompanied by a decrease in its SD; Figure A2-3, Appendix).

In fact, taking the total number of tweets per hour per day as a percentage of the population of the host cities in the pre-, during, and post-Games epochs is revealing. The size of the increase during the Games compared to the pre- and post-Games epochs tracks recency of hosting for four iterations (Figure 34): Rio's increase is greater than London's which is greater than Sydney's which is greater than Atlanta's. Los Angeles does not peak during the Games. Interestingly, although Melbourne is the host from the most distant past, its increase during the Games is comparable to Sydney's. Its activity may have been boosted by hosting football during the Sydney Games. Melbourne also bid to host the 1996 Games, suggesting a persisting interest in the Games by this particular former host. All of these patterns with the exception of Los Angeles are consistent with the possibility of a legacy.

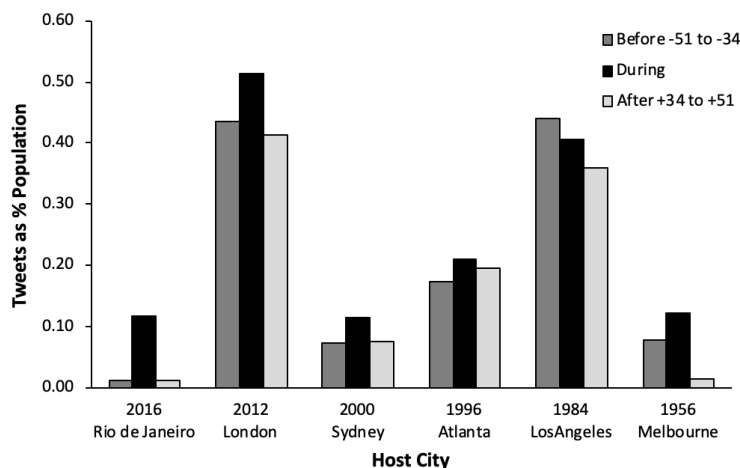


Figure 34. Tweet production as a function of % of population for previous host cities.

DFA

Figure 35 presents results of the DFA for the entire scraping period before the Games, the 17 days of the Games, and the entire scraping period after the Games. Values of H are between .5 and 1.0 for all metrics throughout the scraping period, indicating persistence. However, the values of H seem not to converge as they had in previous the studies.

No previous host city shows the pattern of Rio de Janeiro either for the drop in long-term correlation during the Games for retweets and chatters, or the increase after the Games for tweets, retweets, chatters, and users. In fact, for these Twitter metrics the tendency for the long-term correlation is to drop or not change at all for previous host cities. With respect to the average number of hashtags, the level of dependency peaks during the Games for Rio de Janeiro, Sydney, Los Angeles, and Melbourne. In the case of the average number of words, whereas Rio de Janeiro's coherence peaks during the Games, the previous host cities sustained the long-term correlation patterns of the Games.

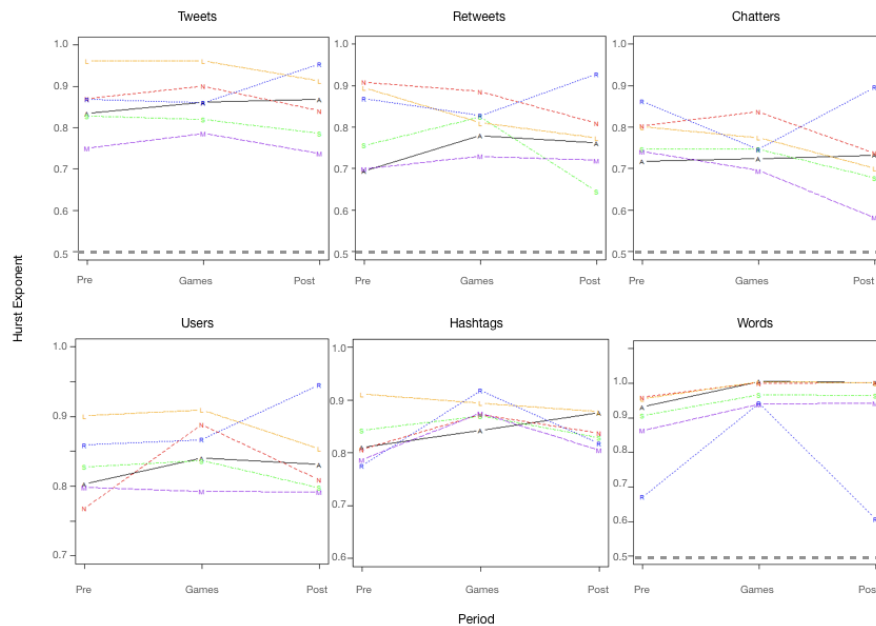


Figure 35. DFA for tweets, retweets, chatters, users, average number of hashtags, and average number of words, pre-, during, and post-Games, in sequence. Cities: Rio de Janeiro (blue dotted line), Atlanta (black line), London (red dashed line), Sydney (green dot-dashed line), Los Angeles (orange two dashed line), Melbourne (purple long dashed line). Dashed grey line indicates the threshold for randomized long-term correlation.

Figure 36 presents the DFA for 9 equal 17-day windows. The results largely echo those for the coarser partitioning. The pattern of long-term correlations observed in Rio de Janeiro (described in Study 2) is not seen in the other former host cities. In fact, for Melbourne, Sydney, and Atlanta there is an increase in dependency during the period of the Games that builds up from some pre-Games windows, and decays window by window, after the end of the event. The DFA does not show a prominent effect of the Games for London and Los Angeles.

The increase in dependency observed in Rio de Janeiro's average number of hashtags is seen in all cities but not in such an exponential manner. As for other Twitter metrics, the growth in long-term correlations is built up in the pre-Games windows. The drop is also not as extreme as the 2016 host city occurring bit by bit in each window.

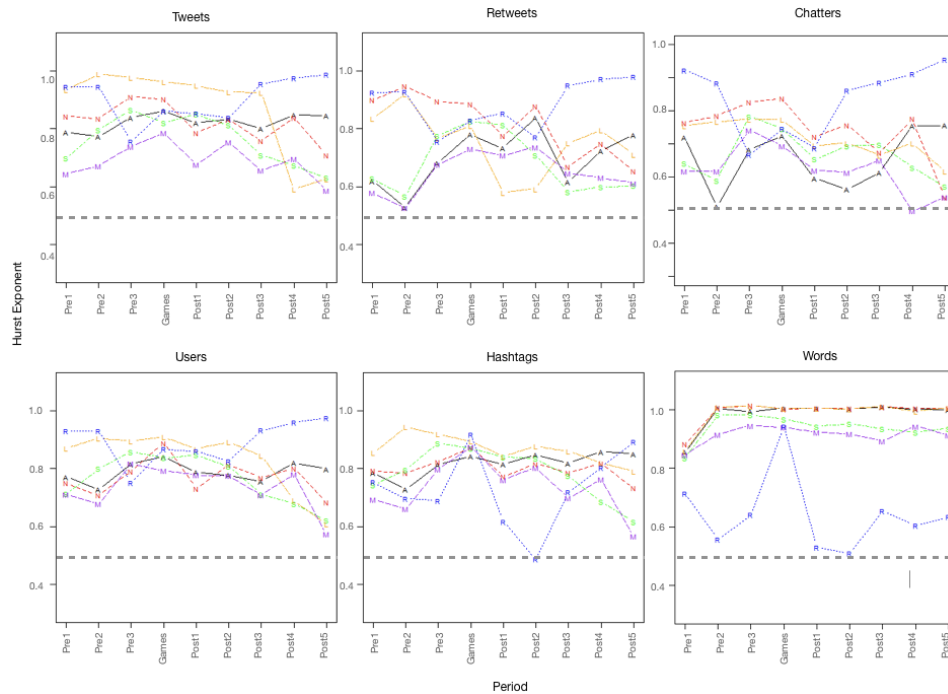


Figure 36. DFA for 17-day windows of tweets, retweets, chatters, users, average number of hashtags, and average number of words per tweet. X-axis in sequence: Three pre-Games windows (starting 51, 34, and 17 days before the Games), Olympic Games window, five post-Games windows (ending 17, 34, 51, 68, and 85 days after the Games). Cities: Rio de Janeiro (blue dotted line), Atlanta (black line), London (red dashed line), Sydney (green dot-dashed line), Los Angeles (orange two dashed line), Melbourne (purple long dashed line). Dashed grey line indicates the threshold for randomized long-term correlation.

Rio de Janeiro also shows a distinct pattern for the average number of words. Its drastic exponential growth in dependency followed by a decay of the same amplitude was not followed by any of the other host cities. Interestingly, all the cities have a very similar long-term correlation tendency that lasts throughout the analyzed windows with not much variation being observed.

These results do not show a clear pattern that is shared by cities that hosted the Games in the past. Nor is it possible to identify any pattern related either to how long ago a city hosted the Games, or to the geographical locality of the city.

FFT

The frequency analysis of the previous host cities echoed the other targeted cities of Studies 1 and 2 in showing the prominence of the 24-hr daily cycle in all windows (Figure 37). The day-night cycle corresponding to a period of 12 hr was again present in the previous host cities. Although it characterized all of Sydney's windows, it appeared in a more mixed fashion for the other cities (sometimes appearing but below the designated threshold). For Los Angeles, the 12-hr period emerged only in windows 5 and 6 (although it was apparent in the other windows, just below threshold). London did not have a significant 12-hr peak in windows 7 and 9, nor Atlanta in windows 1 and 8 (whose magnitudes were below the defined threshold), nor Melbourne in windows 2 and 6. For Rio, it was window 3 that did not have the contribution of the 12-hr period. For those former hosts that showed significant peaks, magnitudes nonetheless varied. In Sydney the 12-hr period had its lower magnitude in windows 2, 3, and 5, localized at the boundaries of the Games. These mixed results for the day-night cycle are of interest because that frequency was prominent in all cities in Morales et al. (2017). Its fragility here could be due to the focus on LTPA, the influence of the Games, the status as former hosts of the Games, or a combination of these.

Slower frequencies than 24 hr were observed in window 1 of London, Sydney, Los Angeles (all 60 hr) and Melbourne (45 hr), and in Sydney's window 5 (90 hr). Faster frequencies than 12 hr were observed in London's window 4 (8 hr) and in Melbourne windows 1 (8 and 2 hr), window 3 (5 and 2 hr), window 8 (10 and 6 hr), and window 9 (7, 3, and 2 hr).

Los Angeles and Atlanta did not have much frequency diversity apart from 24 and 12 hr. Melbourne, in contrast, showed many frequencies in each window (even though some did not quite reach threshold). These results are not clear with respect to the influence of the time elapsed since hosting the Games. The FFT for London and Sydney, the more recent hosts (at 4 and 12 years ago, respectively), showed a bit more variety than Atlanta and Los Angeles (at 20 and 32 years ago, respectively) but not nearly as much as Melbourne the host from 60 years ago. Moreover, Rio de Janeiro, the contemporaneous host, showed diversity during the Games as well as a good deal of below-threshold noise before and, especially, after the Games. In short, the results shown do not seem to reveal any convincing pattern that could be attributed to a city's experience with the Games.

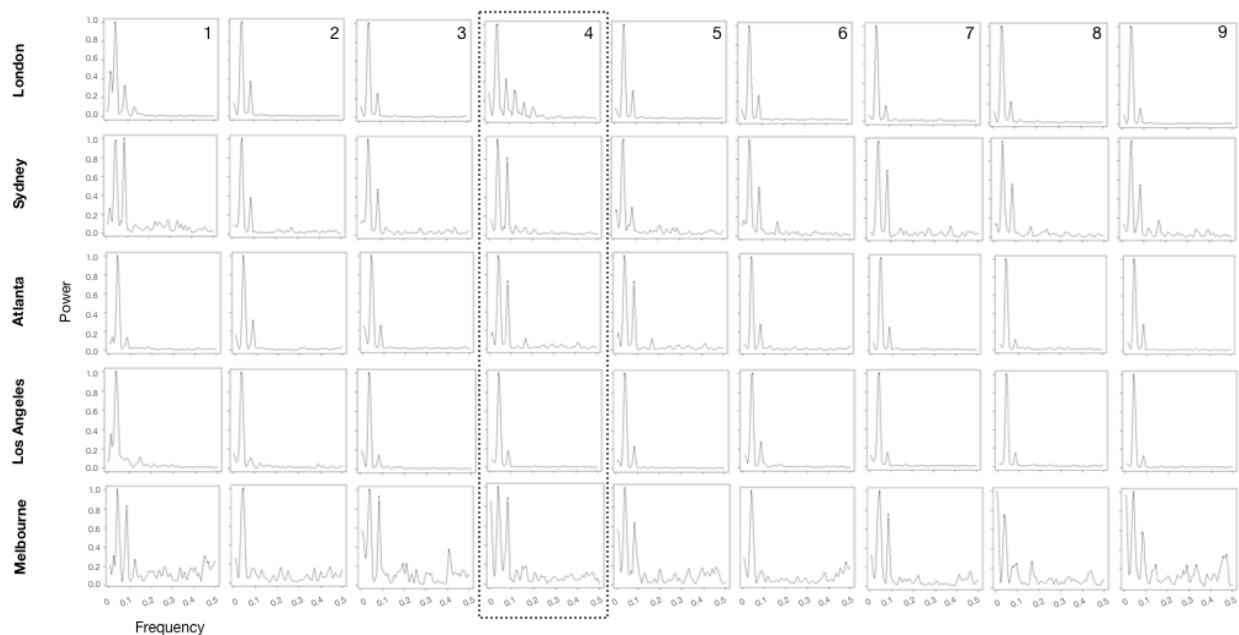


Figure 37. Periodogram plots generated based on the FFT for 17-day windows of tweet activity in London, Sydney, Atlanta, Los Angeles, and Melbourne. (1-3) Pre-Games windows starting at 51, 34, and 17 days

before the Games; (4) Olympic Games window; (5-9) post-Games windows ending at 17, 34, 51, 68, and 85 days after the Games.

Wavelet and Cross-wavelet Coherence

The Wavelet Power Analysis performed on the entire time series for the five former host cities (Figure 38) shows very different patterns from Rio de Janeiro and the other Brazilian cities and more like those of Study 2. The main 24-hr frequency effect evident in Study 2, with a significant high-power period persistent over time, is seen not only in Atlanta and Sydney but also in London, Los Angeles and, less strongly, in Melbourne. And again, unlike the Brazilian cities, the epoch during the Games does not stand out in the spectrum. However, in all the former host cities but Los Angeles, the Games can be detected in the concentration of significant white lines before 2000 hours.

Turning to the finer detail available in the Wavelet Power Analysis performed on the sequence of 17-day windows, the high-power levels around the 24-hr period are sustained over time more than the Brazilian cities, with less persistence in Melbourne (Figure 39). The 12-hr period is less recurrent over time than the 24-hr period, but is visible in all localities, again with less strength in Melbourne.

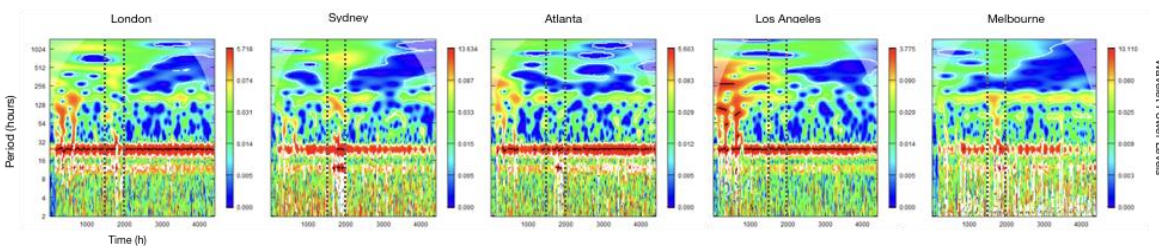


Figure 38. Wavelet Power Spectrum generated for the entire period of tweet activity collected in London, Sydney, Atlanta, Los Angeles, and Melbourne. (Y-axis) Period in hr, (X-axis) Time in hr of the time series. Vertical dashed lines delimit the occurrence of the Games.

In Atlanta, London, and Sydney, the 24-hr period is disrupted during the Games. The impact of the Games is not seen in Los Angeles except for the 12-hr period gaining a little more consistency in terms of increase in power and significance, which also happened in Sydney and Melbourne. However, the opposite effect happened in London, that is, the 12-hr period decreased its significance and power over

time. Another effect of the Games seen in window 4 of all localities, but Los Angeles—the appearance of vertical, high power significant sections of fast frequencies, which extend from 1–24 hr. This shows that many frequencies are responsible for the final time series of tweet production in these former host cities.

One goal of Study 3 was to assess whether being a former Olympic host mattered to the effect of the Games on Twitter activity about LTPA. The described behavior does not seem to be a matter of having received the event in your locality. If that was so, the behavior seen in Rio de Janeiro would be expected to resemble the previous host cities, which did not happen. Save for a few similarities, the frequency spectrum of these five cities is much different from the behavior seen among the Brazilian cities.

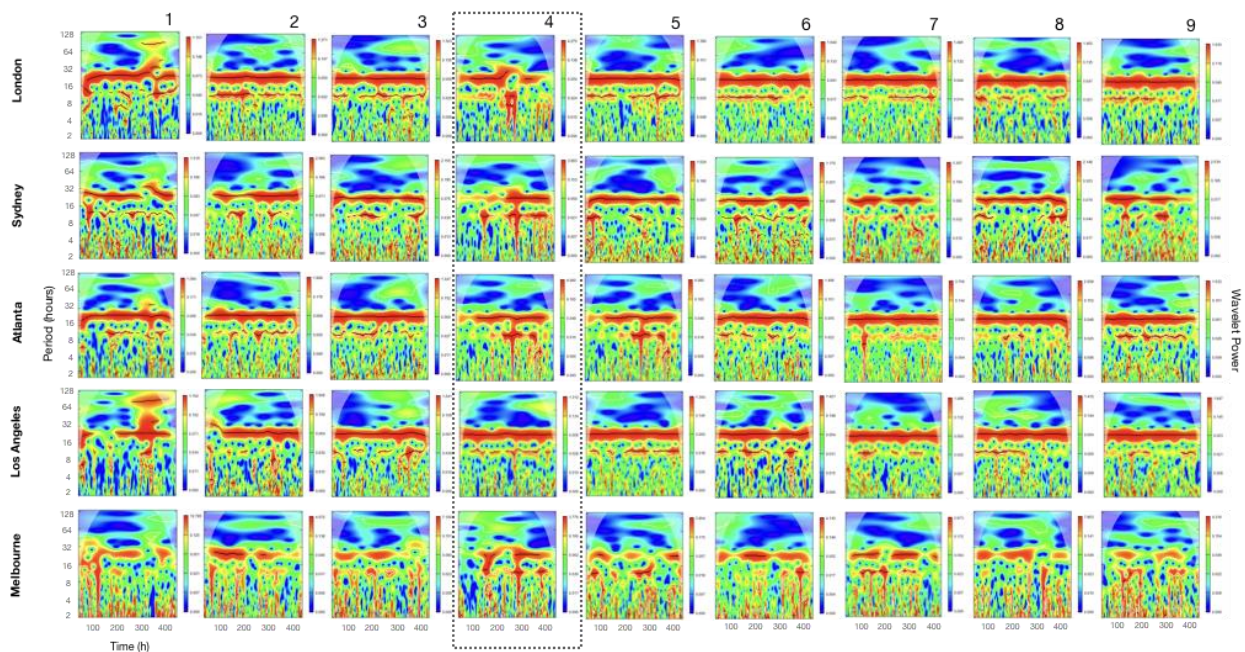


Figure 39. Wavelet Power Spectra generated for 17-day windows of tweet activity in London, Sydney, Atlanta, Los Angeles, and Melbourne. (Y-axis) Period in hours, (X-axis) Time in hours of the time series. (1-3) Pre-Games windows starting at 51, 34, and 17 days before the Games; (4) Olympic Games window; (5-9) post-Games windows ending at 17, 34, 51, 68, and 85 days after the Games).

The frequency coordination analysis performed by means of the cross-wavelet coherence between Rio de Janeiro and the former host cities confirmed the insights allowed by the wavelet analysis. This

can be seen in Figure 40 and in Table 6. The 24-hr period and its surroundings has a high significant correlation throughout the entire analyzed period. The 12-hr period is also prominent, providing evidence of coherence between Rio and the former host cities. The frequency coordination is mostly in phase for these periods. The plots also reveal that all the cities present a disruption in coherence for the period close to the beginning and end of the Games, a likely effect of the event.

Table 6
Cross-Wavelet Coherence of Tweets (with Relative Phase) between Rio De Janeiro and Former Host Cities (with Host Year) as a Function of Epoch Relative to the Olympic Games for Each of Three Periods.

City	Epoch	Period (hr)	% Significance	Relative Phase M (SD)
London (2012)	51-34 days prior	6	20.60	-0.20 (1.68)
		12	51.00	0.09 (0.98)
		24	77.00	-0.38 (0.37)
	17 days during	6	11.03	-0.12 (2.14)
		12	37.80	0.64 (1.02)
		24	63.00	-0.83 (0.50)
	34-51 days after	6	12.50	0.15 (1.93)
		12	66.00	-0.23 (0.66)
		24	98.30	-0.27 (0.23)
Sydney (2000)	51-34 days prior	6	8.33	0.08 (1.68)
		12	54.17	-0.04 (0.93)
		24	52.21	-0.70 (0.72)
	17 days during	6	17.90	-0.07 (1.56)
		12	47.30	-0.22 (0.91)
		24	80.90	-1.54 (0.35)
	34-51 days after	6	11.03	0.48 (1.98)
		12	54.90	-0.95 (0.72)
		24	84.60	-0.52 (0.48)
Atlanta (1996)	51-34 days prior	6	28.20	0.35 (1.66)
		12	44.40	0.01 (1.08)
		24	100.00	-0.46 (0.19)
	17 days during	6	18.14	-0.29 (1.66)
		12	38.00	-0.14 (1.00)
		24	78.70	-0.98 (0.30)
	34-51 days after	6	15.20	-0.18 (1.94)
		12	56.62	0.25 (0.59)
		24	97.31	-0.55 (0.34)
Los Angeles (1984)	51-34 days prior	6	24.30	0.06 (2.00)
		12	34.60	-0.50 (1.43)
		24	75.50	-0.69 (0.23)
	17 days during	6	11.50	-0.17 (1.96)
		12	47.80	-0.77 (1.23)
		24	68.40	-1.22 (0.28)
	34-51 days after	6	9.10	0.02 (1.84)
		12	73.30	-0.51 (0.42)
		24	82.60	-0.47 (0.30)
Melbourne (1956)	51-34 days prior	6	11.03	-0.05 (1.86)
		12	43.40	0.26 (1.20)
		24	58.10	-0.53 (0.44)
	17 days during	6	15.20	0.10 (1.74)
		12	49.80	-0.59 (1.03)
		24	48.80	-1.27 (0.63)
	34-51 days after	6	10.80	-0.24 (1.81)
		12	72.30	-0.10 (0.55)
		24	56.90	-0.59 (0.53)

Los Angeles, London, Sydney and Melbourne present an increased coherence level around a period of 256 hr at the point where the event is happening. At this point, the coordination seems to be shifted towards anti-phase. Both Sydney and Melbourne show a high significant coherence in the slower frequencies around period of 1024 hr, which persists for a great portion of the time analyzed. The frequency coordination tendency at this point is in-phase. The same tendency after the Games is observed in London, but with less coherence strength. This could simply be an indication of a characteristic way these regions tweet.

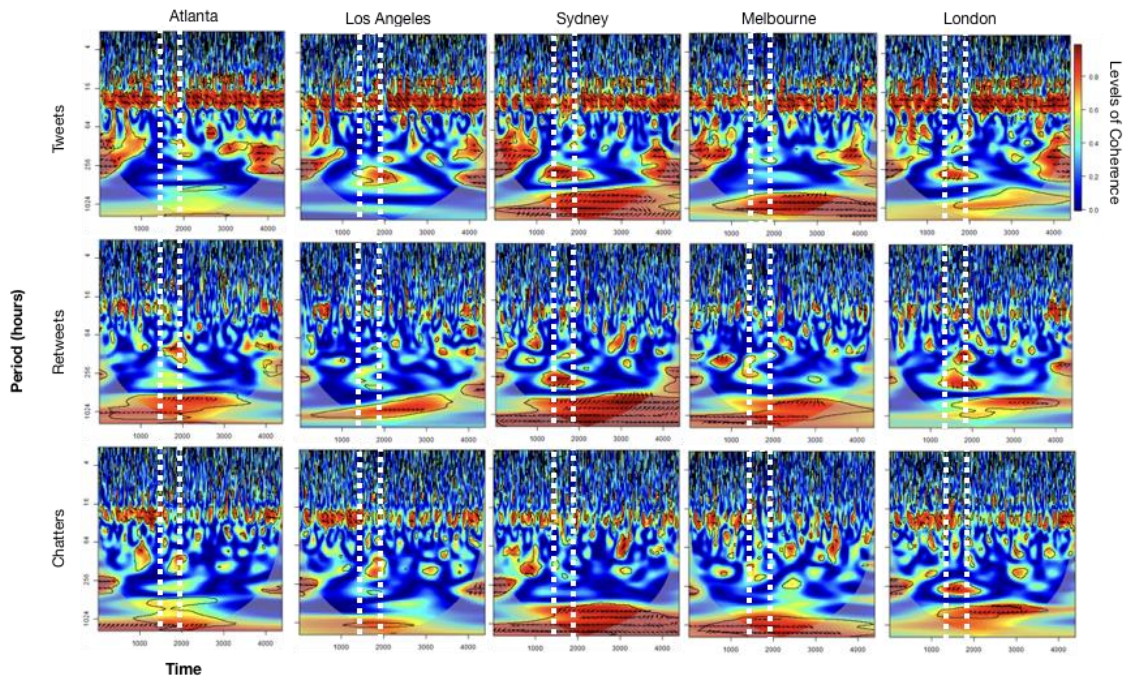


Figure 40. Cross-wavelet Coherence of activity in Rio de Janeiro with (left to right) Atlanta, Los Angeles, Sydney, Melbourne, and London during the entire scraping period for (top) tweets, (middle) retweets, and (bottom) chatters. The Y-axis shows the candidate frequencies; the X-axis shows the time from the beginning to the end of the timeseries (a total of 4391 hr). Vertical dashed white lines delimit the occurrence of the Games (between hour 1560 and 1967).

Summary

The influence of the Games on LTPA Twitter activity was not as clear in the former host cities. Contagious growth is not apparent in the cumulative sums of the Twitter metrics. Nor is there convergence on the degree of persistence during the Games. Moreover, these cities seem to share

neither the promotion of shorter frequencies nor the elevated power of frequencies during the Games that was seen in the Brazilian cities. There is a modest suggestion of a legacy, however, in that Tweet production as a function of % of population is elevated during the Games more for recent host cities.

Selections of cities for particular relationships with Rio de Janeiro produced mixed results. Certain patterns seem to be characteristic of all cities, regardless of functional distance, time zone proximity, or recency of hosting. We now turn to two types of analysis that consider all cities together.

CHAPTER 7

Comparisons of all cities

Cities in Studies 1-3 were selected as a function of their hosting status and their temporal and spatial proximity to the contemporaneous host. It is also fruitful, however, to examine all of the cities together, irrespective of whether or when they hosted the Games. This was done in two types of analyses, one derived from the cumulative sums and the other using Growth Curve Modeling.

Slopes of Cumulative Sums during Three Epochs

As noted in the treatment of cumulative sums in Studies 1-3, cities differ in their general use of social media (e.g., due to population, accessibility of wireless networks, local habits, etc.). Scaling each metric by the total activity during the scraping period, so that they all run from 0 to 1, facilitates comparisons across cities. Figure 11 makes it quite clear that all cities in the host country of Brazil increased their rate of tweeting dramatically during the Games. This is less apparent in Figures 23 and 33 showing the cities of Studies 2 and 3, respectively. Nonetheless, it is possible to discern a subtle shift in the rate of accumulation of tweets during the Games compared to the epochs before and after. This was evaluated by taking the slope of the cumulative sum of tweets/sec for each epoch for each city and treating individual cities as participants in the ANOVA. In order to avoid the possibility that the Brazilian cities would overwhelm the analysis making it appear that the influence of the Games was more pervasive than it really was, the analysis was limited to the seven foreign cities. A one-way ANOVA revealed a significant main effect of epoch, $F(2,18) = 16.98, p < 0.001, \eta^2 = 0.65$. Planned comparisons using Tukey's HSD showed that both pre-Games (Mean = .83) and post-Games (Mean = .79) were significantly different from the Games themselves (Mean = 1.12), $p < 0.001$. This result was replicated in the same analysis limited to the Brazilian cities, $F(2, 6) = 54.65, p < 0.001, \eta^2 = 0.95$. Planned comparisons using Tukey's HSD showed that both pre-Games (Mean = .58) and post-Games (Mean = .49) were significantly different from the Games themselves (Mean = 3.19), $p < 0.001$.

Given that there were three cities in each of three countries—Brazil, Australia, and the U.S—all of the cities with the exception of London can be examined together. A 3×3 mixed design ANOVA revealed main effects of epoch, $F(2, 12) = 60.25, p < 0.0001$, and country, $F(2, 6) = 33.05, p < 0.006$, as well as an interaction of epoch and country, $F(4, 12) = 30.54, p < 0.001$. Figure 41 shows that all countries increased the rate of tweeting during the Games (by Dunn’s Multiple Comparison Test with Bonferroni correction, $p < .0001$ for both Pre-During and Post-During; the Pre-Post difference was ns). Although the average slopes for the Brazilian cities was greater overall ($p < .001$ for both Brazil-Australia and Brazil-U.S.; the Australia-U.S. difference was ns), the significant interaction indicates that this was owing to the steeper boost in the rate of tweeting in the Brazilian cities during the Games.

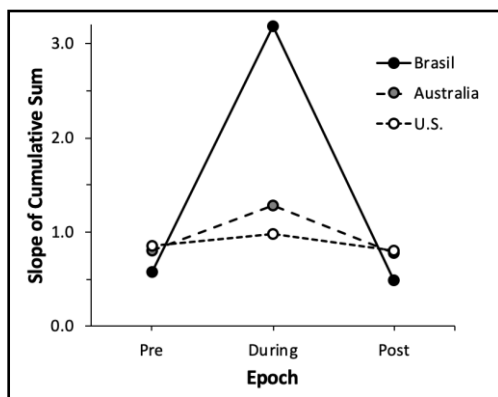


Figure 41. Average slope of the normalized Cumulative Sum of tweets for Brazilian, Australian, and US cities pre-, during, and post-Games.

Growth Curve Modeling

The 10 cities from all 3 studies were included in the Growth Curve Modeling analysis. The first approach was to test the effect of the epochs relative to the Games (pre-, during, and post-Games) as the level 1 time-variant predictor on the variability of the tweet production (normalized). The results indicated that the unconditional model with the epochs as reference had the best basic fit, $\chi^2(7) = 4970.1, p < 0.001$. The residual variance within each city across all epochs was .38 before the Games and .30 after the event. The covariance of the random effects with the epoch of the Games was $r = -0.95$ for pre-Games and $r = -0.97$ for post-Games, indicating that the tweet production characterized by slopes

and intercepts in pre- and post-Games epochs were negatively correlated with the epoch during Games. Pre- and post-Games covariance was high ($r = 0.85$). Fixed effects results indicate that the epoch of the Games accounts for much more change in tweet production, with both pre- and post- epochs decreasing tweet production with respect to the epoch of the event: pre-Games, $t(10.014) = -3.93, p < 0.01$ ($b = -0.776$); post-Games, $t(10.014) = -4.97, p < 0.001$ ($b = -0.865$). In other words, moving from tweet production during the Games to the pre- or post-Games epochs is associated with a decrease in the tweet production.

One issue raised in Study 2 concerned the proximity of a city's time zone relative to Rio de Janeiro. But including this as a predictor did not contribute significantly to the model, $X^2(3) = 4.54, p = 0.209$. Study 3 focused on the effects of being a host city. Including this as a predictor did not contribute to a better fit for the model either, producing a model that did not differ from the unconditional one, $X^2(3) = 0.82, p = 0.85$.

We have seen that cities in the host country altered their rate of tweet production more than the foreign cities. Including status as a Brazilian city as a predictor in the model contributed to a better fit, $X^2(3) = 20.05, p < 0.001$. The total residual variance within each city was 0.89, with the contributions of the pre-Games epoch 0.09 and of the post-Games epoch 0.04. The covariance of the Games random effects with the other epochs was negative (pre-Games: $r = -0.77$ and post-Games: $r = -0.82$) indicating that tweet production during the Games, characterized by slopes and intercepts, was inverse that in pre- and post- epochs. As for the covariance between the pre- and post-Games epoch's random effects, it was positively correlated ($r = 0.26$), meaning the trend in number of tweets followed the same direction of growth. Indeed, the Games changed the fashion of tweet production drastically. Fixed effects result indicated main effects of the epochs with both pre- and post-Games epochs posing a severe tweet production decrease with respect to the Games: pre-Games, $t(9.980) = -3.60, p < 0.01$, ($b = -0.421$); post-

Games, $t(10.123) = -6.38, p < 0.001, (b = -0.535)$. Being a Brazilian city deeply influenced tweet production as well, contributing to an increase in the number of tweets compared with the foreign cities, $t(10.072) = 7.95, p < 0.001 (b = 1.027)$. The model also showed interactions between pre- and post-Games epochs and being a Brazilian city, with these cities having a deeper decay in tweet production at these moments: pre-Games, $t(9.980) = -5.54, p < 0.001 (b = -1.183)$; post-Games, $t(10.123) = -7.17, p < 0.001 (b = -1.098)$.

The final model using epochs as the timescale of reference was:

$$E[TweetsNorm_{it}] = b_0 + b_1Epochs + b_2Brazilian + b_1Epochs*b_2Brazilian + u_{0i} + u_{1i} Epochs \quad (E1)$$

where t indexes time, i indexes individual cities, and the time linear predictor (i.e., Epochs).

The second approach was to test the effect of the epochs relative to the Games with a Piecewise Growth Model in order to segment the time trend into different pieces. The results indicated that the same unconditional model with the time epoch as reference organized in separated dummy coded variables had the best basic fit, $X^2(11) = 4970.1, p < 0.001$. For this model, comparing with the previous unconditional model for the epochs in the same variable, deviance was the same. The total residual variance within each city was lower than 0.001. The covariance of the random effects was effectively zero for the pre-Games relative to the Games ($r = 0.01$), negative for the post-Games ($r = -0.54$), and for the Games with the post-Games epoch ($r = -0.80$). This indicated that the tweet production characterized by slopes and intercepts of the cities during and after the Games tended to opposite directions since they were negatively correlated. Fixed effects results indicated that when compared to the post-Games epoch, the pre-Games epoch did not represent much change in tweet production, $t(9.90) = 0.86, p = 0.411 (b = 0.09)$. However, the epoch of the Olympic Games accounted for such an increase in tweet production that it was significantly different from the subsequent epoch, $t(10.051) =$

4.97, $p < 0.001$ ($b = 0.865$). In other words, moving from tweet production during the Games to the post-Games epoch signified a decrease in tweet production.

Once again, including time zone proximity as a predictor did not really contribute to the model, $X^2(3) = 4.54$, $p = 0.209$, nor did the effect of being a host city, $X^2(3) = 0.82$, $p = 0.85$. And once again, including status as a Brazilian city as a predictor in the model contributed to a better fit, $X^2(3) = 20.05$, $p < 0.01$. The total residual variance within each city was below 0.001. The covariance of the random effects indicated an inversely proportional tendency of increase in tweet production between the Games and the pre-Games ($r = -0.22$) or the post-Games ($r = -0.13$), which reveal the difference in the patterns of tweet production built up during the Games. Once more, the Games are shown to drastically change the style of tweet production. Fixed effects results indicated a main effect only of the Games epoch which presented a severe tweet production increase with respect to the post-Games epoch, $t(10.009) = 6.35$, $p < 0.001$ ($b = 0.535$). The model also showed an interaction of epoch and status as a Brazilian city where the increase in the number of tweets during the epoch of the Olympic Games was greater than compared with the foreign cities, $t(10.009) = 7.14$, $p < 0.001$ ($b = 1.098$).

The final model using epochs as separate timescale variables of reference was:

$$E[TweetsNorm_{ti}] = b_0 + (b_1Epoch1 + b_2Epoch2 + b_3Epoch3) + b_4Brazilian + (b_1Epoch1 + b_2Epoch2 + b_3Epoch3) * b_4Brazilian + u_{0i} + u_{1i} b_1Epoch1 + u_{1i} b_2Epoch2 + u_{1i} b_3Epoch3 \quad (E2)$$

where t indexes time, i indexes individual cities, and the time linear predictors are Epochs 1, 2, and 3.

These two models were using the same time source but organized in different manners inside the analysis and presented the same deviance, even though they had 4 degrees of freedom of difference. No real divergence was detected between them. Also, another aspect that should be highlighted is that

these two approaches compare three epochs of different sizes, corresponding to 65 days before the Games, 17 days during the Games, and 101 days after the Games.

Another aspect of this dataset is that data are not normally distributed in most of the cities except for London and Los Angeles. To try to overcome this fact, the analysis was adapted with the Poisson family, which adds a logarithm to the model. In this third approach the R function used was the glmer, and the dependent variable was the number of tweets without normalization. The unconditional model with the time windows indicated that the epoch as reference had the best basic fit, $X^2(7) = 53657$, $p < 0.001$. The amount of residual variance within each city across all windows was below 0.61 for the pre- and post-Games epochs, and 1.39 for the Games. The covariance of the random effects for the cities' slopes and intercepts with the Olympic Games window was below $r = 0.39$ for both pre- and post-Games. Significant fixed effects were found for pre-Games, $z = -3.85$, $p < 0.001$ ($b = -0.743$), and post-Games, $z = -3.86$, $p < 0.001$ ($b = -0.798$). The result corroborates the positive contribution to LTPA tweet production previously identified during the Games.

Results for cities that were in proximate time zones to Rio de Janeiro indicated that including this predictor did not contribute to the model, $X^2(3) = 6.16$, $p = 0.104$. Including the hosting predictor contributed to a better fit for the model, $X^2(3) = 10.40$, $p < 0.05$, with pre- and post-Games variance below 0.55 each, and a covariance pre- and during Games of $r = 0.24$ and during and post-Games of $r = 0.15$. Significant main fixed effects were found for pre-Games, $z = -4.48$, $p < 0.001$ ($b = -1.03$), post-Games, $z = -4.50$, $p < 0.001$ ($b = -1.103$), host cities, $z = 3.71$, $p < 0.001$ ($b = 1.820$), but no interactions. Including the Brazilian cities predictor contributed to a better fit of the model, $X^2(3) = 45$, $p < 0.01$, with pre- and post-Games random effects variance below 0.03 each, and a covariance pre-Games of $r = 0.87$ and post-Games $r = 0.49$. There were significant main fixed effects for pre-Games, $z = -4.67$, $p < 0.001$ ($b = -0.346$), post-Games, $z = -4.40$, $p < 0.001$ ($b = -0.393$), and host cities, $z = 3.85$, $p < 0.001$ ($b = 1.721$),

but not for Brazilian cities, which was unexpected based on the previous models tested. Interactions were found between status as a Brazilian city in the pre-Games, $z = -11.70, p < 0.001$ ($b = -1.363$), and in the post-Games, $z = -10.90, p < 0.001$ ($b = -1.414$), and between Brazilian cities and the host cities in the pre-Games, $z = -2.38, p < 0.05$ ($b = -0.468$), and in the post-Games, $z = -2.88, p < 0.01$ ($b = -0.569$).

The final model using the Poisson distribution of reference was:

$$\text{Log}(E[\text{Tweets}_{ti}]) = b_0 + b_1\text{Epochs} + b_2\text{Host} + b_3\text{Brazilian} + b_1\text{Epochs} * b_2\text{Host} * b_3\text{Brazilian} + u_{0i} + u_{1i} \quad (\text{E3})$$

where t indexes time, i indexes individual cities, and the time linear predictors are Epochs.

The fourth modeling test aimed to check the effect of the epochs relative to the Games in 17-day windows as in previous analyses. Windows 1-9 constituted the level 1 time-variant predictor of the variability of tweet production (normalized). The unconditional model with the time windows indicated that the epoch as reference had the best basic fit, $X^2(75) = 5323, p < 0.001$. The total amount of residual variance within each city across all windows was 0.88, with window 4 of the Games contributing more (0.21) and windows 8 and 9 above 0.11. The covariance of the random effects for the cities' slopes and intercepts with the Olympic Games window was negative with pre-Games windows 1 and 3, having a more extreme effect for window 1 ($r = -0.76$), and being uncorrelated with window 2. As for the Games relative to the post-Games, the covariance was positive, having higher effects for windows 5 and 6 immediately after the Games (0.49). Significant fixed effects were found for Games window, $t(10.073) = 3.48, p < 0.01$ ($b = 0.556$) and window 9 which is up to 85 days after the Games, $t(10.382) = -2.61, p < 0.05$ ($b = -0.329$). In the case of the Games' epoch, the result corroborates its known positive contribution to tweet production, previously highlighted in this research. The result for the last window of this analysis actually points out that the distance from the Games decreases tweet production about LTPA, a tendency shown in windows 5 to 8, but with only marginal significance levels.

The predictor associated with proximity of time zone to Rio de Janeiro did not really contribute to the model, $\chi^2(11) = 13.90$, $p = 0.238$. Nor did the effect of having hosted the Games, $\chi^2(11) = 17.37$, $p = 0.10$.

Status as a Brazilian city again contributed to a better model fit, $\chi^2[11] = 28.19$, $p < 0.01$. The total residual variance within each city in all windows was $r = 0.88$, with windows 2, 8, and 9 contributing with 0.08, 0.14 and 0.09 variance, respectively. The covariance of the random effects of window 4 with the others was negative for windows 1 ($r = -0.36$), 2 ($r = -0.29$), and 3 ($r = -0.63$). Positive covariance was evident for windows 5 ($r = 0.28$) and 6 ($r = 0.49$), decreasing in the subsequent windows up to negative in window 9 ($r = -0.16$). The fixed effects result revealed a main effect for the epoch of the Games, $t(14.164) = -2.63$, $p < 0.05$ ($b = 0.264$), and windows 5, $t(15.340) = -2.27$, $p < 0.05$ ($b = -0.218$), and 9, $t(10.771) = -3.29$, $p < 0.01$ ($b = -0.462$). The negative slope indicated a tendency for a decrease in tweet postings about LTPA after the Games. The model significant interactions were present in Brazilian cities in the window immediately before the Games, $t(12.195) = 2.29$, $p < 0.05$ ($b = 0.468$), and during the Olympic Games, $t(14.164) = 5.31$, $p < 0.001$ ($b = 0.972$). It demonstrated that, compared with foreign cities, being part of the host country increased tweet production about LTPA.

The final model using 17 days windows was:

$$E[\text{TweetsNorm}_{ti}] = b_0 + b_1 \text{Timewindows} + b_2 \text{Brazilian} + b_1 \text{Timewindows} * b_2 \text{Brazilian} + u_{0i} + u_{1i} \text{Time windows} \quad (\text{E4})$$

where t indexes time, i indexes individual cities, and the linear predictor was Timewindows.

The results of the models make sense since they show that the group of Brazilian cities have a different behavior, especially days immediately before and after the Games as well as during the event. But considering all the options, what would be the best timescale to learn about the changes in this

system? To pursue this question, another proposition was to access the sequential hours as the main time scale (Figure 42), where we can generate a common tendency line for all tweet production in all cities.

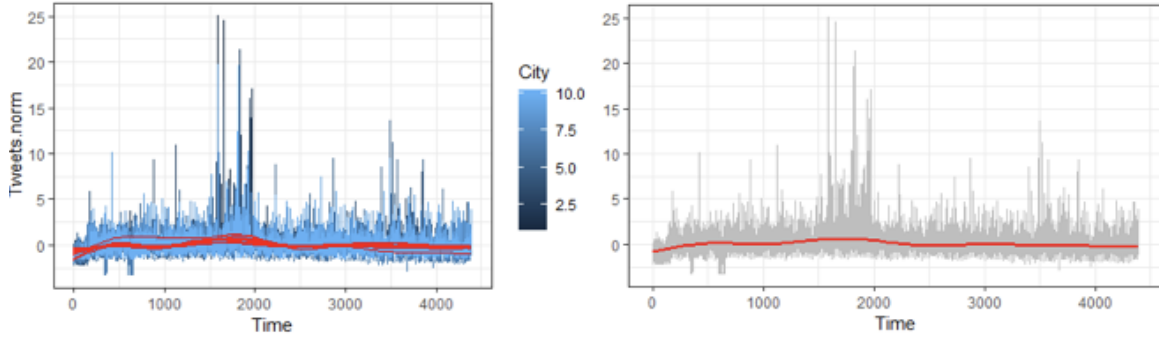


Figure 42. All cities tweet production with their respective tendency lines (left) and the average tendency lines of all cities together (right).

The results indicated that the unconditional model with the time epoch as reference had the best basic fit, $X^2(3) = 1058.7$, $p < 0.001$. The residual variance within each city was close to zero (< 0.001) with a correlation of the random effects of $r = -0.98$. Fixed effects results were not significant. Including time zone close to the Games as a predictor did make a difference in the model, $X^2(2) = 16.74$, $p < 0.001$. In this case, the total residual variance was close to zero (< 0.001), the covariance of the random effects was $r = -0.99$, and there were no main effects or interactions in the model. The time zone predictor was kept as a variable in the expanded model with host cities, but this version of the model was not significantly different from the previous one with time zone only. The same occurred with the inclusion of the Brazilian cities as a predictor. Testing the random effect of epochs on this model resulted in a better fit $X^2(7) = 3885$, $p < 0.001$. The residual variance within each city, of time, and of epochs was close to zero (< 0.001) and a correlation of the Games random effects with Time was $r = -0.80$, with pre-Games $r = -0.28$, and with post-Games $r = -0.46$. Main fixed effects of interaction were not significant, and the final model was:

$$E[TweetsNorm_{it}] = b_0 + b_1Time + b_2TZ + b_1Time * b_2TZ + u_{0i} + u_{1i}Time + u_{2i}Epochs \quad (E5)$$

where t indexes time, i indexes individual cities, and the linear predictors were Time and Time zone.

Using sequential hours as time predictor did not reveal interference from the predictors tested, except for time zone, which was not coherent with previous models' results or with figures 43 and 44.

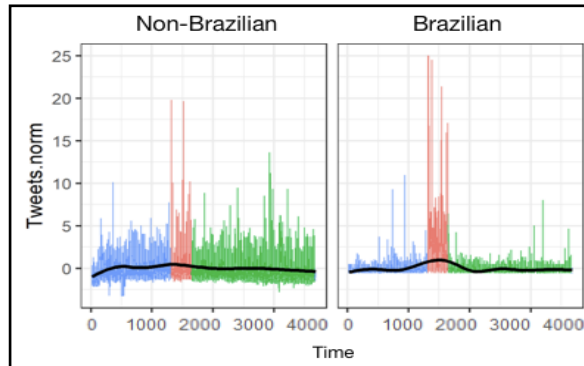


Figure 43. Average tendency line of tweet production for the non-Brazilian cities (left) and for the Brazilian cities (right).

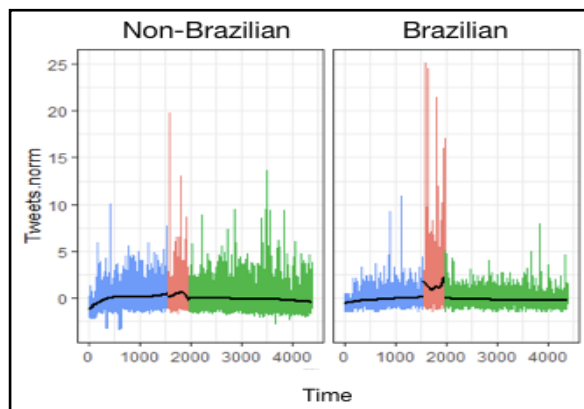


Figure 44. Average tendency line of tweet production for the non-Brazilian cities (left) and for the Brazilian cities (right), with discontinuity between epochs.

Summary

Once again, the influence of the Games on LTPA Twitter activity across all cities is clear. Even though the contagion pattern was not explicit in the non-Brazilian cities, the slope of the increase in the cumulative sums was, in fact, steeper than the slope of the pre- or post-Games epochs for all cities. And GCM again demonstrated that epoch mattered to the production of LTPA tweets. GCM also provided evidence of the importance of functional distance overall in that status as a Brazilian city was a significant predictor of LTPA behavior.

The evidence for a legacy was not found since pre- and post-Games epochs were positively correlated with each other, and never found to be significantly different in the models, meaning that the tweet production at these epochs was statistically the same. The fact that they are also negatively correlated with the Games and were found to be different from this epoch also reinforces the fact that there is not an indication in the GCM analysis that the effect of the Games in tweet production was long-lasting.

CHAPTER 8

General Discussion

Two general aspects of the results were emphasized: (1) Did the Olympic Games influence LTPA Twitter activity; and (2) did the Olympic Games leave a legacy of sports engagement as indexed by LTPA? These aspects were evaluated in cities targeted to represent particular functional and temporal relationships to the Olympic Games. The global patterns of synchrony identified in weekly Twitter activity unrestricted by topic (Morales et al., 2017) provided a benchmark for assessing the patterns in LTPA.

Influence of the Games

From the increase in the rate of tweeting during the Games, especially for cities in the host country, to the change in weekly patterns and increasing variability, the Games had an impact in the targeted cities. This impact is clearly seen in the cumulative sum of tweets among cities in the host country, as well as in the foreign cities, although somewhat less so. In short, and not surprisingly, the Brazilian cities felt the greatest influence from the event but it was not exclusive to them.

The DFA also showed an increase in the long-term dependency of all Brazilian cities during the Games. Moreover, it revealed that the event brought those cities to the same level of dependency, no matter their state in the pre-Games epoch. For the foreign cities, this influence remains—the converging of the cities on particular dependency levels—but in a less consistent fashion.

The wavelet analysis indicates that the Games generally increase, or disrupt, the magnitude of tweet production for 12 and 24 hr cycles. In all three studies, fast frequencies gain magnitude during the event at localized moments reinforcing that times of engagement in tweeting about LTPA get more synchronized during the event throughout the day. The emergence of shorter timescales can be treated

as potential evidence of a social impact of the event, which reflects more interweaving of LTPA in people's lives (suggesting the emergence of objective necessities, at least during the event).

The basic intention of the Cross-wavelet Coherence Analysis is to establish some relationship between the behavior of residents of the host city and the behavior of residents of the other targeted cities. It is an indication of coordination with the particular city that is, in theory, more directly affected by the event. Overall, the % of above-chance coherence for the 24-hr cycle drops during the Games for tweets (and sometimes chatters) for the non-Brazilian cities, breaking a long-term sequence of phase locking, and decreasing the significance of the relative phase which, when present, remains inphase. Mostly, what is seen echoes the DFA and FFT results with respect to the Brazilian cities. The Games seem to increase the power of coherence, making it more consistent over time with long-lasting phase locking, and also increasing the significant role of the faster and slower frequencies inside the Olympic Games timeframe. The patterns are coherent and phase-coordinated during the Games and for the 24-hr cycle. Here also, moreover, the relative phase of coordination is always inphase, meaning that the routine of Rio de Janeiro matches the routine of the other cities at these particular frequencies.

The thesis that the Brazilian cities felt a greater influence of the Games than the other cities was also reinforced by the GCM results. It revealed a greater increase in tweet production during the Games, per epoch, and per window than the other targeted cities. But time zone and host condition predictors were not as consistent in their results, being present in only one of the tested models. Indeed, status as a former host was only manifest in tweet production per epoch, which was higher during the Games for all previous host cities but Los Angeles. As predicted, especially for Rio, the consequences for LTPA were particularly greater due to its involvement in the Olympic Games.

Do the Games Leave a Legacy?

Generally, leaving a legacy is a matter of change emerging in patterns of behavior in economic, social, environmental, and other realms, that resist reversion to earlier patterns once the focal event ends. In this research, the emergent behavior of interest is communication about LTPA. For the contemporaneous host city, for example, heightened LTPA communication could have continued post-Games as an influence from the event. Comparison cities were chosen to see if this influence was stronger for those cities (i) geographically close to the primary host but with different functional roles in the Games, (ii) sharing a time zone with the host city which might simply make it easier to be more attuned to the Games facts and competition, or (iii) former hosts which might be primed to feel the influence of the event due to lasting changes from the past.

From the perspective of the timescale analyses, it is not clear that the Games effected long-term change. By the end of the scraping period, the observed alterations during the Games were already back to their pre-Games states. There was no change that was consistent and persisted until the very last window. Similarly, the longitudinal models did not reveal a robust difference between pre-Games and post-Games. To be sure, the Games differed from the other two epochs. However, the relaxation happened right after the Games, bringing tweet production patterns back to the pre-Games levels. Status as a city in the host country seemed to only play a role during the Games, since pre- and post-Games interactions with this predictor contributed almost the same to the tweet production.

The results of the DFA are more promising, at least for the host city. Rio de Janeiro's tweet production increased its dependency in the time after the Games. This might be an indication of the influence of the Games in establishing social connections, reinforcing the culture of tweeting about LTPA. Since it was the city that experienced the most disruption of its physical environment, for good and ill, this result allows us to think about the potential increase in the practice of LTPA as well, given

evidence of a relationship between the practice of LTPA in a locality and the report of it in Twitter via posts (Liu et al., 2019). The other Brazilian cities do not show a similar change, however, indicating that the functional role with respect to the event might make a difference. But Rio de Janeiro's new routine may be temporary. LTPA Twitter behavior in former host cities does not seem to be related systematically to how long ago they hosted the Games.

The only hint we see of a legacy for former host cities is in how much of a boost LTPA posts receive during the Games. To a limited extent, the size of the boost is related systematically to how long ago a city hosted the Games (Figure 34). But even that pattern breaks down for Los Angeles. The waning legacy of Games hosted in 1984 is not so surprising, of course, given that the age demographic of Twitter-users would not reflect the residents of Los Angeles who experienced those Games first hand. Melbourne, which showed a relatively large increase during the Games, remains a bit of an enigma. Its Games were 63 years ago so clearly the Twitter demographic does not favor direct experience. But as noted, Melbourne has had continued involvement with the Games (bidding to host in 1996, hosting football in 2000), providing opportunities to maintain interest in the Games.

Concluding Remarks

It has been argued that exposure to sports events can lead to increased participation in LTPA (Veal, Tooheyb, & Frawley, 2012; Weed et al., 2009; Weed et al., 2012). Given the substantial investment of financial and human resources required by large-scale sports events, they are a prime topic of discussion (Chen & Henry, 2015; Craig & Bauman, 2014; Feng & Hong, 2013). However, there is a lack of consistency in the literature about whether mega-events do, in fact, play a role in changing behavior related to the practice of physical and sports activities, even in developed countries (Veal et al., 2012; Weed et al., 2009). In fact, considering the difficulties of measurement (Minnaert, 2012; Ritchie et al.,

2009), as well as negative tendencies of social legacies (Kim & Petrick, 2005; Weed et al., 2009), research about the topic has not gotten sufficient attention.

The research reported here addressed this topic through the social medium of Twitter. In particular, this research examined whether emerging dynamic communication about LTPA in the Twitter virtual social environment was influenced by a large-scale event—the Olympic Games—before, during, and after its occurrence. It is important to emphasize that tweets about LTPA are not about the Olympic Games themselves but about individuals' personal physical activity. This distinction is central to the status of LTPA Twitter activity as an index of the practice of physical activities and not Olympics fandom. The main assumption behind this research is that the social dynamics of Twitter usage matches events in the users' physical communities. That is to say, online users talk about what they live in their offline lives (e.g., Quercia, 2012). Of particular relevance for present purposes, they talk about the behavioral possibilities of their environments, especially the ones that they have chosen to actualize.

A formal statement of these possibilities, or affordances, takes the following form:

Situation *X* affords an activity *Y* for organism *Z* on occasion *O* if and only if *X* and *Z* are mutually compatible on dimensions of relevance to *Y*.

Organism *Z* effects an activity *Y* in situation *X* on occasion *O* if and only if *Z* and *X* are mutually compatible on dimensions of relevance to *Y*.

(Turvey, 2019)

The Olympic Games provide the occasion *O* that attunes the Twitter-user *Z* to the affordance of situation *X* for doing activity *Y*. As an example, the situation may simply be *Z*'s neighborhood, which supports the leisure-time physical activity of walking. The occasion of the Olympic Games raises *Z*'s attention to that possibility, prompting *Z* to effect that activity. On the assumption that tweets reflect

actual behavior rather than simply a socially desirable façade (Liu et al., 2019; Quercia et al., 2012), changes in Twitter users' activity provides some evidence that they have become attuned to LTPA-relevant affordances. Importantly, social network activity allows a kind of spontaneity that would not be possible if residents were asked for the information explicitly.

This study stands as an investigation of a perception-action process at the level of the interaction between real and virtual environments—between the Olympic Games and the social network of Twitter—by means of communications about LTPA. As such, it extends the ecological take on action in social contexts. Whatever behavior emerges from a social setting—be it approaching, touching, gesturing, looking, speaking, reading, or tweeting—what is important is the creation of a temporal social bond (Marsh, 2010). Action is treated as a cooperative relational state of organisms embedded within social units (Gibson, 1979; Marsh, 2010). This growing emphasis on the dynamic parallels between intrapersonal and interpersonal coordination (e.g., Richardson, Marsh, Isenhowe, Goodman, & Schmidt, 2007; Schmidt, Bienvenu, Fitzpatrick, & Amazeen, 1998; Schmidt & Fitzpatrick, 2016) suggests that the time is ripe to explore social interaction phenomena more broadly. But it is a challenge to design studies that grasp collective behavior at a larger scale, directly examining many organisms as they interface with each other and with the environment. The effort to understand and examine a complex phenomenon such as self-organizing synchronization is not straightforward, especially when considering large networks (Winfrey, 2002), nested behaviors (Iberall, 1974), and coordination at a global level (Swenson & Turvey, 1991). The study of perception-action by larger collective assemblies that are spatially apart but still connected by sharing the same environment in a regional, continental, or even planetary perspective may provide insight into such complex interactions. The virtual environment, with the kinds of interactions promoted in social networks, is a step in the direction of understanding the issue of collective behavior in the modern world. On a practical level, it may also provide the means for stimulating the practice of LTPA, paradoxically reducing the impacts of the kind of sedentary lifestyle

associated with that virtual environment (Bedimo-Rung, Mowen, & Cohen, 2005; Brasil, 2006, World Health Organization, 2004).

References

- Adnan, M., Leak, A., & Longley, P. (2014). A geocomputational analysis of Twitter activity around different world cities. *Geo-Spatial Information Science*, 17(3), 145–152.
<https://doi.org/10.1080/10095020.2014.941316>
- Aguilera, M., Morer, I., Barandiaran, X., & Bedia, M. (2013). *Quantifying political self-organization in social media. Fractal patterns in the Spanish 15M movement on Twitter*, 395–402.
<https://doi.org/10.7551/978-0-262-31709-2-ch057>
- Roesch, A., & Schmidbauer, H. (2018). WaveletComp: Computational Wavelet Analysis. R package version 1.1. <https://CRAN.R-project.org/package=WaveletComp> A
- Bar-Yam, Y. (1997). The dynamics of complex systems — examples, questions, methods and concepts. In *Dynamics of complex systems* (pp. 01-15). Massachusetts: Addison-Wesley.
- Bedimo-Rung, A. L., Mowen, A. J., & Cohen, D. A. (2005). The significance of parks to physical activity and public health: a conceptual model. *American Journal of Preventive Medicine*, 28(2), 159 –168.
- Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science*, 2(1), 1-8. <https://doi.org/10.1016/j.jocs.2010.12.007>
- BRASIL. Ministério da Saúde. Secretaria de Vigilância em saúde. (2006). *Política Nacional de Promoção da Saúde*. Brasília: Secretaria de Atenção à Saúde.
- Bourdieu, P. (1984). *Distinction: A social critique of the judgement of taste*. Cambridge, Mass: Harvard University Press.
- Brown, C. T., & Liebovitch, L. S. (2010). *Quantitative applications in the social sciences: Fractal analysis*. Thousand Oaks, CA: SAGE Publications, Inc. doi: 10.4135/9781412993876
- Bullough, S. J. (2012). A new look at the latent demand for sport and its potential to deliver a positive legacy for London 2012. *International Journal of Sport Policy and Politics*, 4(1), 39-54.
- Chen, S., & Henry, I. (2015). Evaluating the London 2012 Games’ impact on sport participation in a non-hosting region: a practical application of realist evaluation. *Leisure Studies*, 35(5), 685-707.
<https://doi.org/10.1080/02614367.2015.1040827>
- Craig, C. L., & Bauman, A. E. (2014). The impact of the Vancouver Winter Olympics on population level physical activity and sport participation among Canadian children and adolescents: population based study. *International Journal of Behavioral Nutrition and Physical Activity*, 11(1), 107.
<https://doi.org/10.1186/s12966-014-0107-y>
- Christakis, N. A., & Fowler, J. H. (2010). Social network sensors for early detection of contagious outbreaks. *PLoS ONE*, 5(9), 1–8. <https://doi.org/10.1371/journal.pone.0012948>

- Feng, J., & Hong, F. (2013). The legacy: Did the Beijing Olympic games have a long-term impact on grassroots sport participation in Chinese Townships? *International Journal of the History of Sport*, 30(4), 407–421. <https://doi.org/10.1080/09523367.2013.765725>
- França, U., Sayama, H., Mcswiggen, C., Daneshvar, R., & Bar-Yam, Y. (2016). Visualizing the “heartbeat” of a city with tweets. *Complexity*, 21, 280-287.
- Fusaroli, R., Perlman, M., Mislove, A., Paxton, A., Matlock, T., & Dale, R. (2015). Timescales of massive human entrainment. *PLoS ONE*, 10(4), 1–19. <https://doi.org/10.1371/journal.pone.0122742>
- Gao, C., & Liu, J. (2015). Uncovering spatiotemporal characteristics of human online behaviors during extreme events. *PLoS ONE*, 10(10), 1–14. <https://doi.org/10.1371/journal.pone.0138673>
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.
- Godbey, G. (2008). *Leisure in your life: New Perspectives*, State College, PA: Venture Publishing.
- Groothuis, P. A., & Rotthoff, K. W. (2016). The economic impact and civic pride effects of sports teams and mega-events: Do the public and the professionals agree? *Economic Affairs*, 36(1), 21–32. <https://doi.org/10.1111/ecaf.12156>
- Hall, C., & Hodges, J. (1996). The party’s great, but what about the hangover? The housing and social impacts of mega-events with special reference to the 2000 Sydney Olympics. *Festival Management and Event Tourism*, 4 (1-2), 13-20.
- Iberall, A. S. (1974). On a thermodynamic theory of history. *General Systems Yearbook XIX*, 201-207.
- Kim, S., & Petrick, J. (2005). Residents’ perceptions on impacts of the FIFA2002World Cup: the case of Seoul as a host city. *Tourism Management*, 26, 25–38.
- Kugler, P. N., & Turvey, M. T. (1987). *Information, natural law, and the self-assembly of rhythmic movements*. Hillsdale, NJ: Erlbaum.
- Kulshrestha, J., Kooti, F., Nikraves, A., & Gummadi, K. (2012). Geographic dissection of the Twitter network. *Proceedings of the 6th International AAAI Conference on Weblogs and Social Media*, 202–209. Retrieved from <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM12/paper/view/4685/4984>
- Liu, S., Chan, B., & Kuo, A. (2019). Monitoring physical activity levels using Twitter data: infodemiology study. *Journal of Medical Internet Research*, 21(6). <https://doi.org/10.2196/12394>
- Marsh, K. (2010). Sociality from an ecological, dynamical perspective. *Grounding Sociality: Neurons, Mind, and Culture*, (August), 43–71. Retrieved from <https://doi.org/10.4324/9780203842553>

- Mazo, J. Z.; Rolim, L. H.; Dacosta, L. P. (2008). Em busca de uma definição de legado na perspectiva de megaeventos olímpicos. In L. P. Dacosta et al. (Eds.) *Legado de megaeventos esportivos* (pp. 117-119). Brasília: Ministério do Esporte.
- McBride, J. (2018). The economics of hosting the Olympic games. *Council on Foreign Relations*.
<https://www.cfr.org/backgrounder/economics-hosting-olympic-games>
- Minnaert, L. (2012). An Olympic legacy for all? The non-infrastructure outcomes of the Olympic Games for socially excluded groups (Atlanta 1996-Beijing 2008). *Tourism Management*, 33, 361-370.
- Morales, A. J., Vavilala, V., Benito, R. M., & Bar-Yam, Y. (2017). Global patterns of synchronization in human communications. *Journal of the Royal Society Interface*. Retrieved from
<https://doi.org/10.1098/rsif.2016.1048>
- Pentland, A. S. (2014). *Social physics: How good ideas spread – the lessons from a new science*. New York, NY: The Penguin Press.
- Quercia, D., Diarmuid, O., & Crowcroft, J. (2012). Talk of the city: our tweets, our community happiness. In Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media Talk (pp. 555–558). <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM12/paper/view/4611/5056>
- Richardson, M. J., Marsh, K. L., Isenhowe, R. W., Goodman, J. R. L., & Schmidt, R. C. (2007). Rocking together: Dynamics of intentional and unintentional interpersonal coordination. *Human Movement Science*, 26(6), 867–891. <https://doi.org/10.1016/j.humov.2007.07.002>
- Richardson, M. J., Shockley, K., Fajen, B. R., Riley, M. A., & Turvey, M. T. (2008). ecological psychology: six principles for an embodied-embedded approach to behavior. In P. Calvo & T. Gomila (Eds.), *Handbook of cognitive science: an embodied approach* (pp. 161–189). Amsterdam, The Netherlands: Elsevier.
- Riley, M. A., Richardson, M. J., Shockley, K., & Ramenzoni, V. C. (2011). Interpersonal synergies. *Frontiers in Psychology*, 2(MAR), 1–7. Retrieved from <https://doi.org/10.3389/fpsyg.2011.00038>
- Ritchie, B. W., Shipway, R., & Cleeve, B. (2009). Resident perceptions of mega-sporting events: a non-host city perspective of the 2012 London Olympic Games, *Journal of Sport & Tourism*, 14(2-3), 143-167.
- Romero, V., Fitzpatrick, P., Roulier, S., Duncan, A., Richardson, M. J., & Schmidt, R. C. (2018). Evidence of embodied social competence during conversation in high functioning children with autism spectrum disorder. *PLoS ONE*, 13(3). <https://doi.org/10.1371/journal>.

- Schmidt, R. C., Bienvenu, M., Fitzpatrick, P. A., & Amazeen, P. G. (1998). A comparison of within- and between-person coordination: Coordination breakdowns and coupling strength. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 884-900.
- Schmidt, R. C., Carello, C., & Turvey, M. T. (1990). Phase transitions and critical fluctuations in the visual coordination of rhythmic movements between people. *Journal of Experimental Psychology: Human Perception and Performance*, 16(2), 227–247. Retrieved from <https://doi.org/10.1037/0096-1523.16.2.227>
- Schmidt, R. C., & Fitzpatrick, P. A. (2016). The origin of the ideas of interpersonal synchrony and synergies. In P. Passos, K. Davids & C. J. Yi (Eds.), *Interpersonal coordination and performance in social systems* (pp. 17-31). London: Routledge.
- Speed, B. (2014). Which city tweets the most? CityMetric. Retrieved from <https://www.citymetric.com/which-city-tweets-most>
- Sporting Goods Manufacturers Association (2012). SGMA says the Olympics do impact sports participation. Sports & Fitness Industry Association. Retrieved from <https://www.sfia.org/press/431>
- Statista (2018). Most popular sports events during the 2015/2016 season, sorted by number of tweets (in millions). Retrieved from <https://www.statista.com/statistics/290997/popular-sports-events-tweets-united-states/>
- Swenson, R., & Turvey, M. T. (1991). Thermodynamics reasons for perception-action cycles. *Ecological Psychology*, 3(4), 317-348.
- Swinnen, J., & Vandemoortele, T. (2008). Sports and development: an economic perspective on the impact of the 2010 World Cup in South Africa. *ICSSPE Bulletin*, v.53, p.1-6, Amsterdam: Drukkerij Libertas, Bunnik.
- Tan, S., Li, Y., Sun, H., Guan, Z., Yan, X., Bu, J., ... He, X. (2014). Interpreting the public sentiment variations on Twitter. *IEEE Transactions on Knowledge and Data Engineering*, 26(5), 1158–1170. Retrieved from <https://doi.org/10.1109/TKDE.2013.116>
- Gouhier, T. C., Grinsted, A., & Simko, V. (2018). R package #> biwavelet: Conduct Univariate and Bivariate Wavelet Analyses #> (Version 0.20.17). Available from #> <https://github.com/tgouhier/biwavelet>
- Thelwall, M., Buckley, K., & Paltoglou, G. (2011). Sentiment in Twitter events. *Journal of the American Society for Information Science and Technology*, 62(2):406-418.
- Turvey, M. T. (2019). *Lectures on perception: an ecological perspective*. New York: Routledge.

- Veal, A.J., Tooheyb, K., & Frawley, S. (2012). The sport participation legacy of the Sydney 2000 Olympic Games and other international sporting events hosted in Australia. *Journal of Policy Research in Tourism, Leisure & Events*, 4(2), 155–184.
- Xu, Y., Liu, Z., Zhao, J., & Su, C. (2017). Weibo sentiments and stock return: A time-frequency view. *PLoS ONE*, 12(7), 1–21. <https://doi.org/10.1371/journal.pone.0180723>
- Weed, M., Coren, E., Fiore, J., Wellard, I., Mansfield, L., Chatziefstathiou, D., & Dowse, S. A. (2009). *Systematic review of the evidence base for developing a physical activity and health legacy from the London 2012 Olympic and Paralympic Games*. Centre for Sport, Physical Education & Activity Research (SPEAR)/Canterbury Christ Church University.
- Weed, M., Coren, E., Fiore, J., Wellard, I., Mansfield, L., Chatziefstathiou, D., & Dowse, S. (2012). Developing a physical activity legacy from the London 2012 Olympic and Paralympic Games: a policy-led systematic review. *Perspectives in Public Health*, 132(2), 75-80.
- Winfrey, A. (2002). Oscillating systems. On emerging coherence. *Science*, 298(5602), 2336-7.
- World Health Organization (2004). “Estratégia Global em Alimentação Saudável, Atividade Física e Saúde”. (Tradução não oficial realizada pela CGPAN/MS). 57ª assembléia mundial de saúde, WHA.

Appendix 1

The Twitter metrics

The Twitter metrics analyzed were the number of tweets, retweets, chatters, users, the average number of hashtags, and the average number of words. The tweets are the regular posts limited to 140 characters. Retweets are tweets that were forwarded to a user's group of followers. The chatters are tweets that were addressing an user's name which is done by @ followed by the user's name. The number of users is the total amount of accounts that posted something during the hour analyzed. The hashtags correspond to the main topics to which users classified their tweets and is accompanied by the symbol #. The average number of hashtags corresponds to the average amount of topics that were showing in the tweets produced per hour. The average number of words corresponds to the average amount of words used by users in their tweets aver an hour.

Filters

Identifying LTPA hashtags

A set of search terms or hashtags related to LTPA was prepared (i.e., words or expressions that highlight and categorize the accompanying text). These terms, which were written in English and Portuguese, were defined through intense manual exploration of the Twitter database. In a snowball approach, each hashtag candidate (e.g., fitness, sport), was manually typed on the platform and any associated hashtags found in the tweets that appear in the search were included in the list until no new hashtags were identified. During the process, all identified hashtags included in the list were typed in the platform. Hashtags that appeared in only one tweet accompanied by other popular hashtags already listed were not selected. A total of 163 hashtags were identified. In addition, hashtags that accompanied other topics unrelated to LTPA (e.g., fashion, music) or used in order to produce the opposite meaning (e.g., not), were used as filters to exclude posts form the search. A range between 25,000 and 160,000 tweets per day was obtained with this technique.

Identifying localities associated with targeted cities

In order to gather tweets from the targeted cities, considering that each locality is manually entered by each user who writes their city names in a variety of ways, a list of the names of the localities found in the network was used. The names of the localities grouped as the same city were selected by inspecting lists of localities. After that, scripts in the R programming language were created for grouping targeted localities used in the studies in separate data frames that contained the necessary variables for running analyses. First, a search in Google was performed to uncover possible names a city could have. For example, a search about Rio de Janeiro city (as distinct from Rio de Janeiro state) returned certain terms like rio, Rio, RIO, and RJ. All the possible names and spellings of the city name that were found were included in the R script, narrowing it to around 5000 locality terms. Each result was inspected manually to guarantee it was related to the city of Rio de Janeiro and not another place. If any doubt persisted about the connection of the locality term with the targeted city, the term was excluded from the analysis.

Appendix 2

Detailed Patterns

Study 1

Considering the average number of words per tweet, it is apparent that Belo Horizonte has the highest sum and mean values, followed by Rio de Janeiro then Porto Alegre. The maximum number of tweets undergoes oscillations among all windows. While Rio de Janeiro and Porto Alegre's maximum number of words increases 17 days before the Games, followed by a drop during the Games. For Belo Horizonte, the increase happens 34 days pre-Games, followed by a drop, and another increase during the Games. In the case of the standard deviation, Porto Alegre presents higher values followed by Rio de Janeiro and Belo Horizonte, which present closer statistic values. The drop in standard deviation observed during the Games in these cities is immediately followed by a rise in variability after in the first window after games that is maintained in the following windows (Figure A-1).

Study 2

The exceptions hold for Rio de Janeiro in most of the Olympic Games window, for Sydney on sum and mean statistics of the average number of words per tweets per hour. Atlanta's metrics lead in almost all windows of hashtag activity. Atlanta's maximum values and standard deviation also became the lowest of all targeted cities for all the analyzed windows of the category average number of words per tweets per hour, showing that the routine of posts is basically the same for all the periods analyzed and does not vary with the sports event.

Several metrics (tweets, retweets, chatters, and users) increased during the Games for all targeted cities (Figure A-2). It is apparent that the effect of the Olympic Games pushes Rio's metrics beyond the other targeted localities, except for average number of hashtags and words. Atlanta produces greater values especially for sum and mean statistics of tweet activity and in most of the windows, with a few exceptions. The variability is constrained by the event 34 days before it begins and remains this way until 85 days after it ends in a consistent manner for Atlanta, Sydney and Philadelphia. Brisbane returns to its pre-Games first window statistic values post-Games. Looking at the patterns for hashtags, the increase occurs pre-Games close to the beginning of the event in most of the cases, with a drop in the first window post-Games, especially sharp for Rio de Janeiro, showing that the effect of the event in the posts about LTPA on the network lasts mostly the time of the Games.

Study 3

Twitter metrics descriptives are displayed in three following figures corresponding to the groups of cities of each study. Some of it, is explored in the body of text, mostly focusing on tweets, retweets, and chatters.

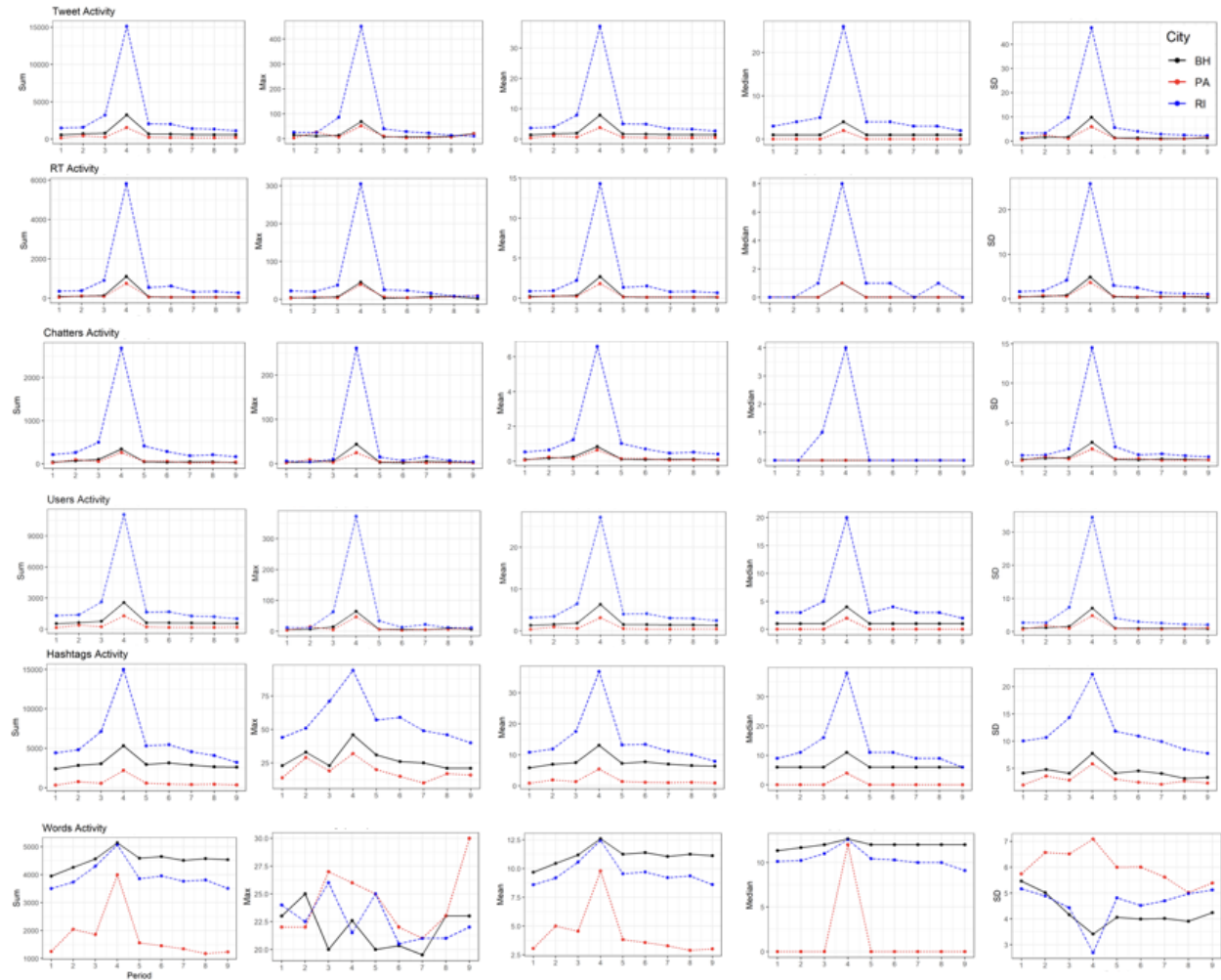


Figure A-1. Study 1's Descriptive Statistics (sum, maximum number, mean, median, and standard deviation) for tweets, retweets, chatters, users, average number of hashtags, and average number of words. X-axis: (1-3) Three pre-Games windows (starting 51, 34, and 17 days before the Games); (4) Olympic Games window; (5-9) five post-Games windows (ending 17, 34, 51, 68, and 85 days after the Games). Rio de Janeiro (blue dotted line), Belo Horizonte (black line), Porto Alegre (red dashed line).

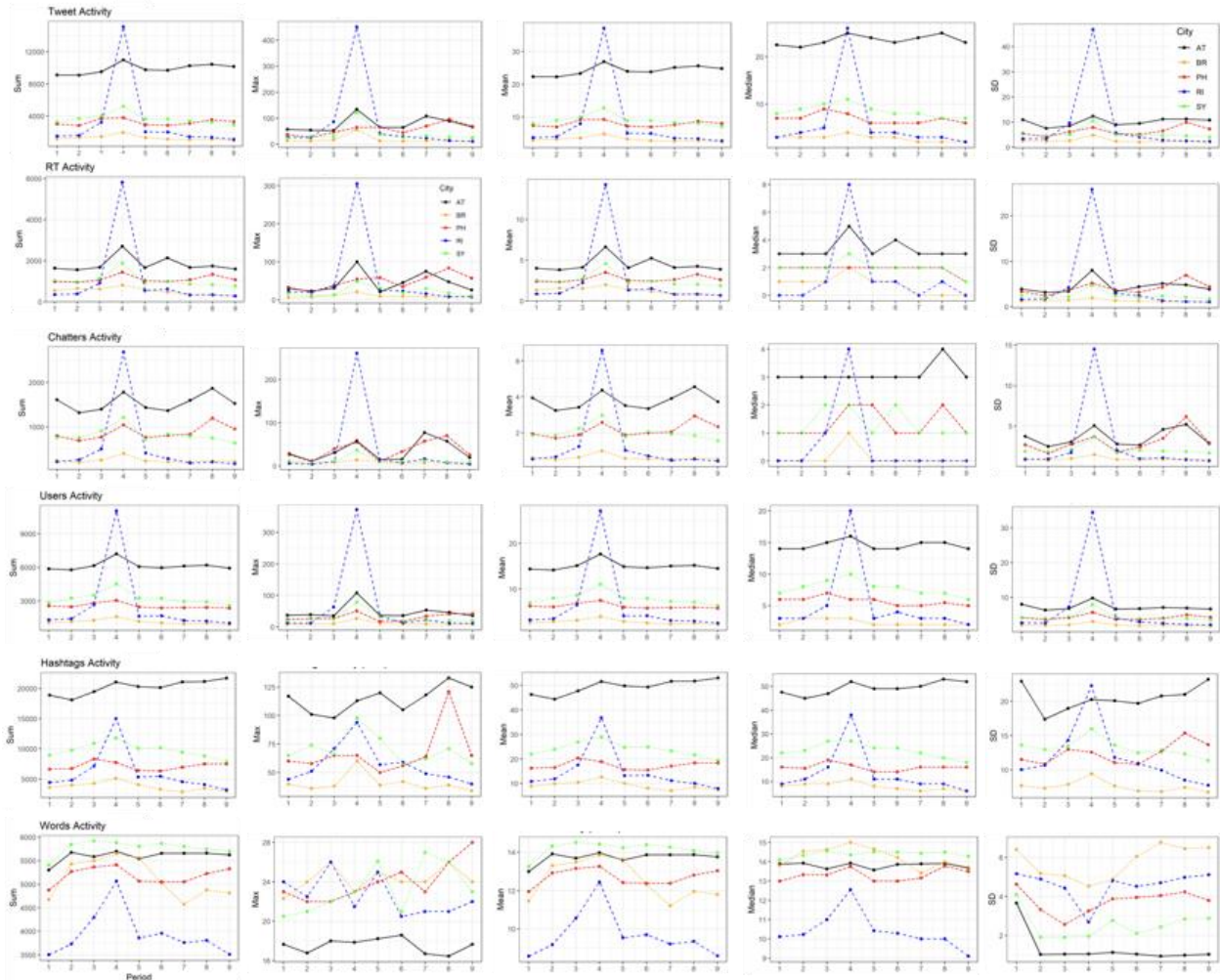


Figure A-2. Study 2's Descriptive Statistics (sum, maximum number, mean, median, and standard deviation) for tweets, retweets, chatters, users, average number of hashtags, and average number of words. X-axis: (1-3) Three pre-Games windows (starting 51, 34, and 17 days before the Games); (4) Olympic Games window; (5-9) five post-Games windows (ending 17, 34, 51, 68, and 85 days after the Games). Rio de Janeiro (blue dotted line), Atlanta (black line), Philadelphia (red dashed line), Sydney (green dot-dashed line), Brisbane (orange two dashed line).

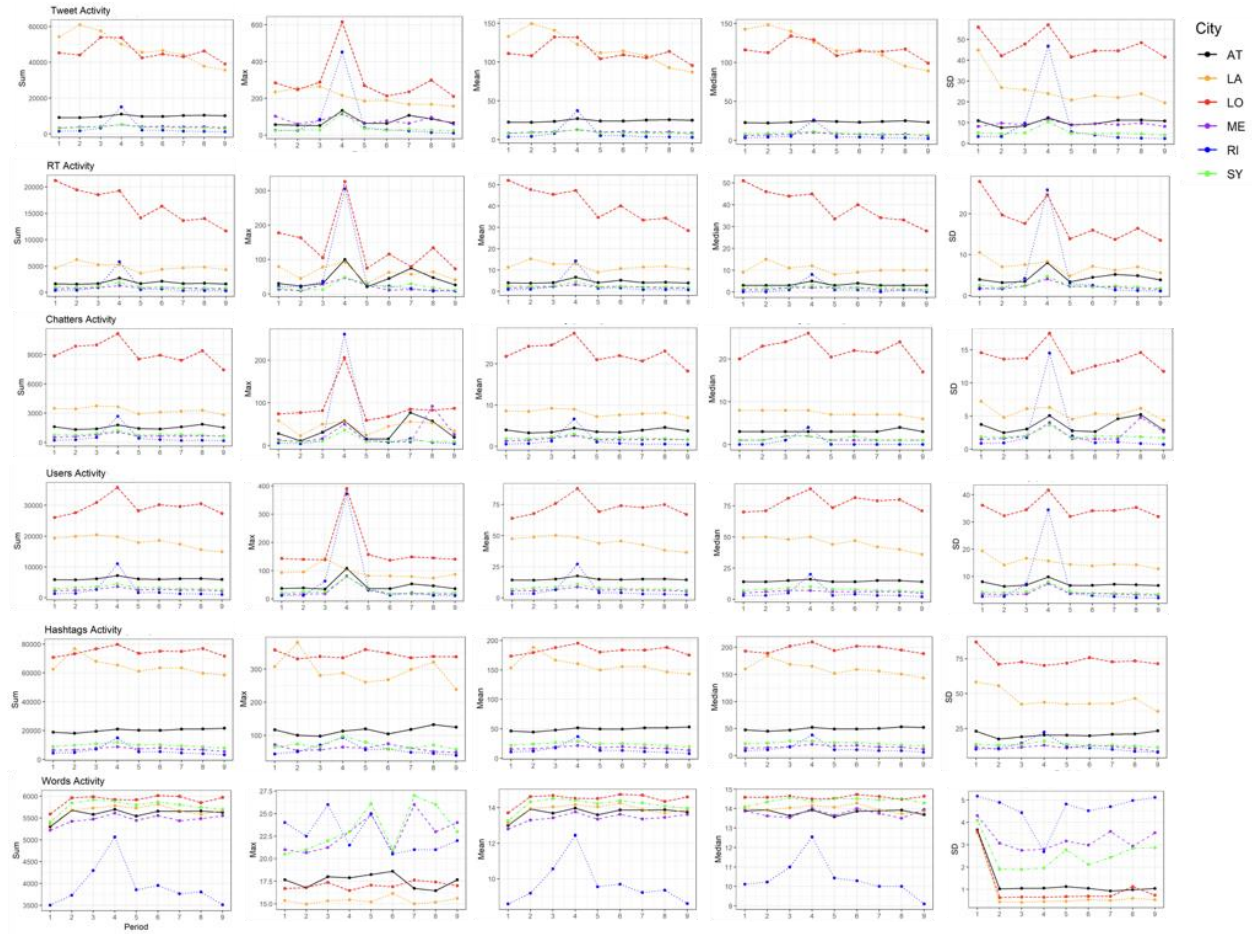


Figure A-3. Study 3's Descriptive Statistics (sum, maximum number, mean, median, and standard deviation) for tweets, retweets, chatters, users, average number of hashtags, and average number of words. X-axis: (1-3) Three pre-Games windows (starting 51, 34, and 17 days before the Games); (4) Olympic Games window; (5-9) five post-Games windows (ending 17, 34, 51, 68, and 85 days after the Games). Rio de Janeiro (blue dotted line), Atlanta (black line), London (red dashed line), Sydney (green dot-dashed line), Los Angeles (orange two dashed line), Melbourne (purple long dashed line).